DOCUMENT RESUME

ED 050 021 24 SP 004 850

AUTHOR Burdin, Joel L., Ed.; And Others
TITLE Accreditation and Evaluation of Basic Teacher Education Programs: Research Problems and Prospects.
INSTITUTION American Educational Research Association, Washington, D.C.; ERIC Clearinghouse on Teacher Education, Washington, D.C.
SPONS AGENCY National Center for Educational Communication (DHEW/OE), Washington, D.C.
BUREAU NO BR-8-0490
PUB DATE May 71
CONTRACT OEC-0-8-080490-3706(010)
NOTE 70p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Accreditation (Institutions), Program Evaluation, *Research Needs, *Standards, *Teacher Education, Teacher Evaluation

ABSTRACT

The five papers collected in this document were delivered at the 1970 American Educational Research Association symposium sponsored by the Special Interest Group on Teacher Preparation Curriculum. All five focus on developing a research base for teacher education standards, in particular for the "Recommended Standards for Teacher Education" by the American Association of Colleges for Teacher Education. John Herbert advocates accreditation criteria based on teacher behaviors and different standards for different types of programs. To build the required knowledge base, he would research, in part, the cut-off points and standards applied in practice by accrediting teams. S. C. T. Clarke summarizes curriculum standards and notes that the emphasis placed on certain standards rather than the standards themselves will cause change. In the third paper, Donald M. Medley offers ways of closing the research-practice gap, among them a centralized information exchange and a one-shot questionnaire which would solicit information on teacher education program characteristics and on the success of beginning teachers. John R. Dettre focuses on the current state of literature and research in the areas of admission, selective retention, and student involvement in program development and evaluation. The final paper, by R. L. R. Overing, is a review of studies which evaluate teacher education graduates according to behavioral objectives and pupil change. Bibliographic items total 211. (LP)

Brief Title:

ACCREDITATION AND EVALUATION
RESEARCH PROBLEMS

ED050021

ACCREDITATION AND EVALUATION OF BASIC TEACHER EDUCATION PROGRAMS: RESEARCH PROBLEMS AND PROSPECTS

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Edited by Joel L. Burdin
and Margaret T. Reagan
ERIC Clearinghouse on Teacher Education

Compiled by John Herbert
The Ontario Institute for Studies in Education
Toronto, Ontario

Published by
ERIC Clearinghouse on Teacher Education
Washington, D.C.

and the

Special Interest Group on the Teacher Preparation Curriculum,
American Educational Research Association

May 1971

SP 004 850

SP004850

TABLE OF CONTENTS

Foreword. ii

About ERIC. iii

A RESEARCH BASE FOR THE ACCREDITATION OF TEACHER PREPARATION PROGRAMS . 1

THE RESEARCH BASE FOR TEACHER PREPARATION CURRICULUM STANDARDS. 18

THE ACCREDITATION RESEARCH BASE RELATIVE TO FACILITIES. 23

STANDARDS OF STUDENT SELECTION AND INTERACTION. 30

RESEARCH RELATED TO THE EVALUATION OF GRADUATES OF TEACHER PREPARATION PROGRAMS 55

FOREWORD

While it is commonplace to read opinions about performance-based teacher education, analyses and interpretations of research are in short supply. These papers, delivered at a symposium sponsored by the Special Interest Groups on the Teacher Preparation Curriculum at an annual meeting of the American Educational Research Association, meet some of the needs for ideas and information on an important topic.

Compiling, editing, and publishing this document does not constitute acceptance of, or support for, the data and analyses reported herein by the Clearinghouse staff. Our role has been one of facilitation of study, and where appropriate to teacher educators, further research and implementation. Likewise this publication does not necessarily reflect endorsement of performance-based teacher education by our sponsoring groups or by the U.S. Office of Education.

Credit for making this publication possible is due the AERA presenters, and especially to Dr. John Herbert, Ontario Institute for Studies in Education, whose interest and persistence were indispensable. Miss Margaret Reagan is to be commended for carrying out many of the tedious but necessary tasks of preparing a publication. Miss Christine Pazak converted manuscripts--marked by Miss Reagan and me--into clean, typewritten copy.

A recently released publication entitled "Performance-Based Certification of School Personnel" is a good companion piece to this one. Compiled by personnel in the Florida Department of Education as proceedings of a conference largely funded by the U.S. Office of Education, the book was prepared for publication by this Clearinghouse and published by the Association of Teacher Educators, one of our sponsors. (Joel L. Burdin and Margaret T. Reagan, editors. Washington, D.C.: Association of Teacher Educators and ERIC Clearinghouse on Teacher Education, 1971. 140p.)

We believe that this publication is an interim one. Much has been done, and is being done, to add specifications to the most important task of preparing school personnel. Varied efforts will produce more knowledge about this topic. ERIC will publicize the contents and availability of the new knowledge base, hopefully stimulating others to broaden and deepen it.

Joel L. Burdin
Director

May 1971

About ERIC

The Educational Resources Information Center (ERIC) forms a nationwide information system established by the U.S. Office of Education, designed to serve and advance American education. Its basic objective is to provide ideas and information on significant current documents (e.g., research reports, articles, theoretical papers, program descriptions, published or unpublished conference papers, newsletters, and curriculum guides or studies) and to publicize the availability of such documents. Central ERIC is the term given to the function of the U.S. Office of Education, which provides policy, coordination, training, funds, and general services to the 20 clearinghouses in the information system. Each clearinghouse focuses its activities on a separate subject-matter area; acquires, evaluates, abstracts, and indexes documents; processes many significant documents into the ERIC system; and publicizes available ideas and information to the education community through its own publications, those of Central ERIC, and other educational media.

TEACHER EDUCATION AND ERIC

The ERIC Clearinghouse on Teacher Education, established June 20, 1968, is sponsored by three professional groups--the American Association of Colleges for Teacher Education (fiscal agent); the Association of Teacher Educators, a national affiliate of the National Education Association, and National Commission on Teacher Education and Professional Standards of NEA. It is located at One Dupont Circle, Washington, D.C. 20036.

SCOPE OF CLEARINGHOUSE ACTIVITIES

Users of this guide are encouraged to send to the ERIC Clearinghouse on Teacher Education documents related to its scope, a statement of which follows:

> *The Clearinghouse is responsible for research reports, curriculum descriptions, theoretical papers, addresses, and other materials relative to the preparation of school personnel (nursery, elementary, secondary, and supporting school personnel); the preparation and development of teacher educators; and the profession of teaching. The scope includes the preparation and continuing development of all instructional personnel, their functions and roles. While the major interest of the Clearinghouse is professional preparation and practice in America, it also is interested in international aspects of the field.*

The scope also guides the Clearinghouse's Advisory and Policy Council and staff in decision-making relative to the commissioning of monographs, bibliographies, and directories. The scope is a flexible guide in the idea and information needs of those concerned with pre- and inservice preparation of school personnel and the profession of teaching.

A RESEARCH BASE FOR THE ACCREDITATION OF TEACHER PREPARATION PROGRAMS*

John Herbert
The Ontario Institute for Studies in Education
and The University of Toronto

The other papers of this symposium ask what light research has thrown and can throw on the present criteria, procedures, and standards of accreditation of basic teacher preparation programs adopted by the American Association of Colleges for Teacher Education Evaluative Criteria Study Committee (and subsequently adopted in 1970 as the official standards of the National Council for the Accreditation of Teacher Education). It is the purpose of this paper to consider how a research base might be established for the development of alternative or supplementary accreditation standards. Such research would deal with questions of curriculum evaluation and design, and with the evidence we have and need in guiding institutions in strengthening their teacher preparation programs.

THE RECOMMENDED STANDARDS

The direction here recommended is in keeping with the current policy of the AACTE as expressed in the new *Recommended Standards for Teacher Education*. While previous drafts were aimed to "help to protect children and youth from ill-prepared school personnel" (3: 1), the newest document clearly states that the goal is to set up procedures which will assure the public that accredited programs "meet national standards of quality," that "children and youth are served by well-prepared personnel," and that the teaching profession is advanced "through the improvement of preparation programs" (4: 1). While the earlier goal merely made it necessary to identify the bad eggs, the new aims require a much clearer knowledge than we now have of the possible meaning of the word "standards" in the current document, the relationship between the nature of programs and the teaching ability of their graduates and the values which should inform efforts toward improving programs. The changes here proposed are also in accordance with some of the most advanced proposals for changes in teacher preparation (56). Before turning to the question of alternative criteria and research evidence, I should like to clarify my position by examining some underlying issues raised by the *Recommended Standards*.

Criteria, Standards, and Values

In discussing these issues it will be helpful to make a distinction between two words which are often used synonymously: "criteria" and "standards." I will use the word "criterion" to refer to a characteristic which is to be examined by an accrediting team. I will reserve the word "standard" for a qualitative or quantitive measure of the degree or extent

---

*A report prepared for the Special Interest Group on the Teacher Preparation Curriculum and Division B, American Educational Research Association, Minneapolis, 1970.

to which a program possesses that characteristic. For example, the following statement in section G4.1 of the *Recommended Standards* is by my definition a criterion: "Standard: the library is adequate to support the instruction, research, and services pertinent to each teacher preparation program" (4: 11). The subsections, which consist of questions pertaining to such matters as diversity of holdings, library use, and annual expenditure, make this criterion more specific, so that the staff of the program knows what characteristics of the library the accrediting team will consider. But only if a minimum standard is explicitly stated (e.g., the library shall contain at least $200 worth of books per student), can the program see how far it must improve its library to attain accreditation. And only if a continuum or set of continua is presented --indicating various standards below and above the minimum for each specific criterion--can a program compare its resources with those of other programs or aim at a given degree of improvement.

With very rare exceptions, the present *Recommended Standards* indicates the general areas of a program to be assessed--that is, it establishes criteria--but it does not state standards. One exception is the requirement that at least one-third of any program must be in liberal studies. Even here, however, it is unclear whether there is also a top limit which in turn constitutes a minimum for other components. It may be argued that the liberal studies component is limited by the requirements of the professional component, but this is described as intended to "provide a set of categories through which an institution can describe and review the professional studies component of the various teacher education curricula it offers." So the requirement here is merely that each program will contain something recognizable as belonging to each named category.

This amounts to a set of criteria analogous to those established for the evaluation of libraries. Such criteria might be thought of as minimum standards: A program must show some evidence of attention to each of the criteria in all the categories in the document. If this is the case, the standards are so low that they are unlikely to serve as incentives for improvement and will at best duplicate state and regional accreditation. More likely, however, there are hidden standards behind the criteria; or each accrediting team must in practice establish its own standards, adjusting them perhaps to the professional goals and values of particular programs it is responsible for examining.

The AACTE has thus tacitly recognized two serious issues: (a) the problem of conflicting values within or between accreditation groups and between accreditation groups and the program to be accredited; (b) the difficulty of obtaining evidence adequate to establish or support standards. Accreditation, even when it is intended merely to ensure minimum resources or continued progress towards any goals a program may set for itself, nevertheless assumes that some values exist. These values may be economic: maximum allocation or optimum use of facilities or of human and curricular resources. They may be pedagogical: preparing the type of teacher held to be most effective in certain kinds of schools. They may simply be developmental: assuring that programs continuously evaluate and modify their own practices. In any case the determination of such values precedes accreditation procedures, even the

establishment of criteria. To list criteria but to leave standards inexplicit does not in itself resolve the problem of conflicting or unsubstantiated values.

Cut-off Points

Research based on the current *Recommended Standards* might help to clarify accreditation procedures and their underlying values by finding out what standards accreditation teams apply in practice, where the cut-off points actually are, how they vary for different types of programs, and how much agreement or disagreement arises within and between accrediting teams and program staffs. We also could seek to determine whether teacher preparation programs or their staffs show any detectable changes when these minimum standards are applied (9).

It seems impossible, however, that the results of such research would not go very far to implement the stated goals of the present accreditation procedures. If the minimum standards applied by accreditation teams are low, then only inferior programs will be affected. If the cut-off points are high, then marginal programs may seek to meet them; but programs which are clearly below or above the standard are unlikely to be much affected. A procedure other than the establishing of a cut-off point seems to be required to establish standards beyond the minimum level which simply ensures that an institution has the resources and facilities necessary to operate a teacher preparation program at all.

Single Versus Multiple Standards

Obviously, if standards are to exceed such a minimum level, there will be a conflict between the desire to apply a single set of standards to all institutions and so "ensure national standards of quality" (4), and the desire to leave institutions free to design their own programs. As has been pointed out, "the single standard necessitates the framing of component criteria in very broad terms in order that they be operable," and as a result they are "often only statements of good intention . . ." (19: 61).

Again, research may help. It would be possible to describe existing programs of various types in terms of the component criteria provided in the *Recommended Standards* and to formulate sets of explicit standards for each component of all the main types of program. The components of every program selected for study could then be rated unsatisfactory, minimal, good, or excellent by standards applicable to programs of that type. The resulting descriptions and ratings, presented as anonymous case studies, could be matched with those of almost any program to be accredited and a corresponding profile identified. As new programs are devised, appropriate descriptions and standards could be added. Instead of a cut-off point, this procedure would employ a series of descriptive-evaluative statements of a number of key components of teacher education programs. It would, moreover, avoid the need to value one educational philosophy over another.

Self-Evaluation

Multiple standards developed in this way would provide a means for detailed assessment and comparison of the facilities, organization, and curricular and human resources of preparation programs. These are the components dealt with in four of the five sections of the *Recommended Standards*. The fifth section, "Evaluation, Program Review, and Planning," also provides a basis for describing resources of a somewhat different kind: The institution must conduct "a well-defined plan for evaluating teachers it prepares," and it must use the evaluation results "in the study, development, and improvement of its teacher education programs." The criteria are the existence and employment of evaluation procedures and plans for modification of programs. The institution and not the accrediting agency is to evaluate the effectiveness of its program and of its efforts at improvement.

These criteria are valuable in themselves insofar as they indicate a necessary component of preparation programs. Self-evaluation is a continuing process while accreditation is an occasional procedure. In stressing self-evaluation, moreover, the *Recommended Standards* is in effect giving credit to programs which have conducted or supported research on teaching behavior. As I shall try to show later, there are several ways in which research can help to implement this recommendation.

It is difficult to see how, finally, an accreditation agency can develop standards for assessing the adequacy of self-evaluation procedures without at some stage making some judgment of the impact of the total program on the teaching behavior of its graduates and on what goes on in the schools. The emphasis on resources is a tradition which goes back at least to the Flowers Report of 1948; but, it has been pointed out, although it was good in its day, "Excellence demands more vigorous research in the future, particularly on the results the programs achieve" (38: 7).

What Should Be Researched?

The question then arises in what way the results of teacher preparation can be assessed--and what kinds of results should be selected for examination. It has often been argued that the validity of the evaluation of a teacher preparation program increases if the evidence is collected as close as possible to the "final product" or the "third level"--the changes in the pupil (62: 62). Though indisputable in theory, this argument does not work in practice. While we should do more and better research on which teacher behaviors result in changes in pupil behavior, it is not expedient to evaluate teacher preparation programs by such changes in the schools where the teachers find employment.

Pupil changes occur to a great number of different individuals, each of unknown personality, unpredictable cultural conditioning, and idiosyncratic response. The reaction to any teacher cannot necessarily be attributed to the teacher and much less to the teacher's preparation. Moreover, pupil changes, except responses to tests, are extremely difficult to record accurately. In any case, such changes occur in environments where the teachers of teachers control only one of the variables

--the training the teacher receives. There is evidence that any effect of the training can go driven underground at least temporarily by the anxieties inherent in beginning teaching.

Combinations of variables--the school and home environment of the pupils, the decisions of the teacher's peers and administrators, and those of the teacher himself--may result in placing him in a position where, regardless of the training received or the criteria used, he either cannot fail or cannot succeed. It would thus be no more reasonable to evaluate a teacher preparation program by the way pupils learn in the classroom of graduates than to evaluate a program of medical training by the health of the population its graduates serve. Therefore, though it is theoretically attractive to relate pupil behavior to accreditation, this seems unlikely to be feasible in the foreseeable future. As Ryans found: "With all the attractiveness of judgment of teacher behavior from its products [e.g. pupil changes] . . . the disadvantages of such approaches seem to outweigh their advantages" (47: 71).

When we concentrate instead on teaching behavior, the chances of obtaining meaningful information become much greater. The available research is growing rapidly. It is already having an impact on teacher preparation (11: 415). We can use the results of direct observation of teaching and also data about indirect variables which may be related to the teachers' preparation. Both are potentially very fruitful lines of evidence for the accreditation of programs and for the improvement of teacher preparation.

New Complementary Criteria

The criteria which I should like to propose differ from those included in the *Recommended Standards* in that they are based on the description not of programs, but of the behavior of teachers by the programs to be accredited. To establish standards based on these criteria would not require the establishment of an invariant relation of particular components of the preparation programs to the behavior of its graduates, since similar results may be produced by disparate causes. Moreover, a program which has achieved results held to be acceptable or desirable should, in the absence of strong counter-evidence, be presumed to employ appropriate means. Programs seeking to improve their effectiveness could act on the information gathered during accreditation by attempting to determine specific factors within or beyond their institutions which might affect the teaching behavior of their graduates in desirable and undesirable ways.

Evidence on the behavior of teachers may be gathered by examining records; by obtaining testimony from students, graduates, or supervisors in oral or written form; and by observation of teaching. There are indications that at least some of these types of evidence discriminate among teacher preparation programs (53; 7), but that each is subject to some limitations which need to be taken into account in formulating criteria and standards. Three of the most promising types of evidence seem to be career line data, client satisfaction, and direct evidence about teaching. Although only a few studies can be cited here, much more research has been done in each area and there are obvious practical reasons why we should consider each area carefully.

Career Line Data

Information about career lines includes such matters as wastage from teaching, ratings and recommendations of supervisors, types of teaching and administrative positions held, participation in research and program development, further training and education undertaken, and so on. Some career line information--for example, wastage from teaching --clearly would be highly relevant and important for accreditation. The present criteria do not call for any information on the number of years that graduates spend in teaching, only on whether or not they enter the teaching profession (4). It might, however, be considered that an average of at least three years teaching is necessary to justify the expenditure of resources in teacher training.

After all, a teacher who remains in teaching for four years costs half as much to educate as two teachers who stay in the profession for two years each. To make wastage a criterion would implement the goals of accreditation since reduction of wastage would cut down the number of inexperienced teachers in the schools and the number of student in preparation programs, thus making for more stable teaching staffs and releasing resources for the improvement of programs.

Information on wastage is already being collected in some states (15; 43) and also by some teacher preparation programs in their follow-up studies reported in the educational journals. Studies of teacher mobility often omit information on training programs in preference to information on age and sex (41), which of course are clues to the incidence of marriage and pregnancy, two major reasons why teachers drop out, but which are probably of less significance to the profession. Researchers may be able to describe groups or types of graduates and the conditions under which they have high or low survival rates. They could investigate variables which seem likely to be related to wastage, including the appropriateness of the new teachers' skills for the initial teaching position, and attempt to locate particular program components or variables which might be altered to reduce wastage. Such research would be of considerable practical and theoretical interest.

Other kinds of career lines data, which from a commonsense position seem to be of at least equal significance, are more difficult to assess than wastage. Ratings by supervisors and peers are an example. Information of this kind is relatively accessible, since it can be gathered directly by interview or other techniques, collected from records, or inferred from positions of appointive or elective responsibility, held by a teacher.

Such information, however, is subject to a number of limitations. Procedures and criteria for evaluating teachers vary from district to district, and frequently the evidence on which ratings are based is very meager or second hand (41). The same problems apply to reports of changes in teachers and teacher growth (59).

The personality of the principal also seems to have a substantial effect on the ratings of a teacher's ability and social competence (61; 60; 21). In addition, school district and college supervisors do not

agree in their ratings of teachers (52). Perhaps the attempt to divide teachers into types based on profiles of attributes they have in the principal's judgment may be more valid (27), but the evidence is not strong.

Ratings by pupils seem to be a much more promising source of information (44). Unfortunately, however, there is evidence that they do not agree with ratings by supervisors (55). This could be a problem from a practical point of view, making it awkward to collect the information in the schools and to explain the results. Again, there are differences in the ratings of teachers by different groups. For example, older students may put more emphasis on scholarship (8).

In view of these limitations and difficulties, it does not seem possible at present to establish workable criteria based on ratings by supervisors, peers, and pupils. Other career line data are still less promising because of a lack of research studies or the inconclusiveness of the evidence so far obtained. The relationship of a teacher's participation in research and program development to his teaching appears to be unstudied. Such variables as experience, competence in the subject field, training in the teaching of that subject, further education, and so on, have been found to be related to teaching competence in some research studies but unrelated in others. Thus Blosser and Howe (8), reviewing 20 studies, find personal adjustment and academic preparation to be related to success in teaching high school science. However, Metzner (39) find on reviewing 17 research studies and reviews of research that there is no evidence of a relationship between the length of a teacher's training and his knowledge of his subject, and supervisors' ratings or pupil achievement, however measured. But administrators apparently believe that teachers should be more specifically trained for particular skills or levels of teaching--especially in stimulating thinking (51). And there is some evidence that those who have completed teacher preparation programs are rated as better teachers than those who have not, even when it may be more suitable for accreditation procedures than most kinds of career line data. The opinions of students and graduates are a good source of subjective evidence about the effects of a program and of its various components. A program whose students hold it in high esteem is probably in a better position to affect their teaching than a program held in low esteem. Such a program is also more likely to be able to obtain information and advice from its graduates when it seeks to evaluate and improve its offerings (33).

It should be noted that the arguments for using testimony of students and graduates do not necessarily hold for other groups--administrators, parents, and educational critics--whose opinions cannot be considered direct evidence of program effectiveness. Such indirect testimony may, however, have implications for program planners--for example, the evidence that most parents associate unpleasant discipline experiences with women teachers and more positive experiences with men teachers (32). For strong practical reasons, too, programs cannot afford to ignore client opinion, since it affects such matters as funding, recruitment of students, and placement of graduates.

Unfortunately, a number of problems make it difficult to conduct valid studies of client satisfaction. Students who are still in a program or who have just begun to teach are not yet in a good position to evaluate its usefulness for teaching. After some years of experience, teachers do not recall the details of their training; and their testimony is harder to collect. Furthermore, students' reactions to a program can vary greatly from year to year, though when these changes are in response to program changes they may be important evidence (26). The attitudes of any single group of students also seem to change during the period of their training (20), probably towards accepting the views of the teacher preparation staff and especially those of the supervising teacher (8). However, the direction of change seems to reverse itself when students graduate and begin professional teaching (12; 54), making it difficult to know when to measure client satisfaction unless these changes prove to be predictable. Graduates may however also be affected by the climate of opinion prevailing at the time of a study, as is suggested by the changes in opinions expressed about methods courses (1; 13). Follow-up studies have had great variation in success in getting responses, varying from 40 percent to over 90 percent, with most around the middle of this range. There is evidence to suggest that the "lost" part of the population differs from the respondents, except when the rate of response is very high (52; 27).

Most follow-up studies are conducted by teacher preparation staff members, who don't have the time, the skill, or the motive to conduct rigorous analyses of the research design, procedures, and results. The problem of bias during the collection, and especially during the analysis of data, is high when the researcher also teaches in or administers the program under study. Of course, such research can be conducted by organizations other than the teacher preparation programs themselves, as was done by the National Union of Teachers in England (42).

The most satisfactory base for accreditation would be a profile of the graduates' teaching derived from a set of measures of their teaching performance in a variety of appropriate situations. There are a number of practical and theoretical reasons why it is essential to include some assessment of the teaching of the graduates of a program in accreditation procedures.

Changes in teacher behavior are obviously the central goal of teacher preparation. Any program that has no detectable impact on its graduates could hardly be considered effective. At the same time, information about how graduates teach is most valuable for the design and evaluation of a program by its staff; and--if carried out carefully and periodically--it would provide a baseline for measuring the impact of subsequent changes in the format, resources, or other variables of the program.

The development of criteria to assess the teaching of behavior of graduates can be also justified on theoretical grounds. The analysis of teaching and of educative relationships is one of the most promising fields of educational research. The number of instruments and techniques is growing rapidly, and our knowledge is increasing both quantitatively and qualitatively. Two very useful anthologies, *Mirrors for Behavior*,

Parts One and Two (49; 50), give information on approximately 80 direct observation techniques, most of them developed quite recently. Without further research, unfortunately, these promising new instruments and techniques for describing, predicting, and evaluating teaching cannot be used for accreditation purposes, since they are still in the development stage. With the goal of accreditation criteria clearly in view, however, research efforts might become better coordinated and more effective.

Each type of evaluation technique has advantages and disadvantages for developing the kind of profile needed. For ease of administration and interpretation the ideal would be a test or battery of tests, with descriptive, value-free norms standardized for different populations. Work is now in progress to develop tests of this kind (37; 23). Unfortunately no test with descriptive or predictive power for teaching has yet been developed.

Another possibility is the use of an easily staged simulated teaching environment which would make it possible to immerse the new teacher in a variety of teaching situations. Work is in progress on such situation tests--using sound film to simulate teaching sequences with facility to change the events simulated (48; 30)--and on microteaching and mini-lessons which provide scale models or analogues of classroom teaching (35; 29). In this work, however, the training effect is often given more emphasis than evaluation.

While each of these procedures has an iconic relationship to actual teaching, they are isomorphic only to a limited extent. Even when they are fully developed, it seems likely that a number of these situation tests would have to be combined to form a battery before one could expect much descriptive or predictive accuracy.

An alternative procedure would be to observe graduates in actual teaching situations in classrooms, in laboratories, and on field trips. Until recently such observations were inevitably unsystematic, and they therefore could not be used to provide precise or objective descriptions of teaching. However, as observation techniques--for example, those collected in *Mirrors for Behavior*--become more fully developed, this drawback should cease to be a problem. Particularly if the sample of teaching behavior is recorded, the material can be collected from a selected sample according to a previously arranged schedule and can be re-examined when evaluations disagree. Analysis of the material by means of a framework developed for the purpose can be done as needed, and one of a set of suitable standards can be applied to categorize the graduates and the program.

The practical problems in the way of direct observation are much less difficult than might be anticipated (25). The theoretical problems are more serious, but these also can be resolved. Techniques of observation and analysis of teaching will have to be standardized. It would not be possible to standardize students or classroom situations, but a rough categorization of teaching situations would probably be adequate. The diversity of possible ways of teaching could make it difficult to establish a profile, but research evidence suggests that teachers actually employ a fairly limited repertoire of teaching styles (6; 22).

This is not the place to review the now extensive literature on the observation of teaching and learning. I believe, however, that we are now developing techniques which, with further research, will become sufficient for describing the teaching of graduates of teacher preparation programs for purposes of accreditation. It seems to me, moreover, that teaching behavior is the ultimate criterion against which all other measures of the effectiveness of programs must in time be validated. The importance of the goal, I believe, should outweigh any other consideration in determining the direction of our research efforts.

Research Base

I have suggested that in establishing a research base for the accreditation of teacher preparation programs we should develop the present criteria into a set of multiple standards to fit diverse programs and several levels of quality. I have also suggested that additional standards for measuring the effects of preparation programs should be formulated. Standards serving this purpose could be based on such criteria as career lines (especially retention in teaching), client satisfaction, and--above all--the teaching behavior of students during the program and after graduation.

In preparing the present paper, I gathered hundreds of papers directly or indirectly relevant to this topic. Many were discussions of criteria for accreditation or proposals for new programs of teacher preparation, often very thoughtful and ably presented; but strangely, even in the best of these discussions there were few if any references to research. At best the authors referred to or had themselves conducted surveys of opinions and attitudes (56; 2). Many authors deplored the lack of research studies and spoke of teacher preparation as an "unstudied problem." Yet despite the paucity of references and the frequent call for more research, there is in fact a very large number of studies that can be drawn upon to inform discussions of accreditation and teacher preparation. Research on the description of teaching and on preparation programs dates back at least to the Commonwealth Teacher Training Study (15). Since that time it has increased greatly, especially in the last decade (31; 7; 3; 24; 14). Can we then draw upon these studies to form a research base for the new criteria? We can, but there are some major problems.

Inadequacy of Reporting

Most of the research studies are reported in journals. A substantial number of other studies remains unpublished (though where these are doctoral dissertations they can be traced). Another large group, especially reports of a follow-up studies, were never completed. Even in the published reports of research, however, much information about the procedures and the milieu of the study is usually omitted. For example, the reports rarely state the type of school or schools in which the research was conducted or describe the educational program of the schools, the socio-economic level of the pupils, the ages and academic and professional preparation of the teachers, the teaching styles they employed, and so on--except when these were the variables directly under study. Yet these

variables clearly often affect the results. A researcher who wanted to make use of a study for almost any purpose would need at the very least to go back to the primary research report. Perhaps he would even need to contact the investigators before he could interpret the results.

The lack of replication is also a very serious problem. Replications of research studies are as rare as reports of peace in the newspapers. Strangely, researchers will often produce a single instance as though it were generalizable. Their next piece of research is usually quite different. No one verifies any results, so that there is no evidence that another experimenter or the same experimenter at another time or place would have obtained the same results. The erratic distribution of research topics is still another problem. Dussault (16) reports that he found 16 studies of the effect of supervision on the attitudes of student teachers, and only one study (10) of the effect of supervision on their teaching behavior. This situation is quite widespread, with the result that no research at all has been done in some key areas.

Heavy reliance on theories external to pedagogy is another problem. Ideally, research studies can be placed into a network of theory which relates them to one another. The lack of a common theoretical framework means that research results often are not directly comparable. Also, it is very difficult to conduct a rigorous review of research, since it is necessary to analyze the terms used and check them against the theory in which they are imbedded in order to interpret the methodology of a study and its results.

Lack of Screening

Perhaps this difficulty is the main reason why these studies are rarely examined and tested rigorously, along the lines of some recent correspondence (46; 57; 45; 58). In the absence of such uniform procedures as are found in the natural sciences, the likelihood of error is very high. The resulting errors may hide significant results or produce a deceptive significance. If enough studies are conducted, the probability of some statistically significant results occuring at random is high. As studies with statistically significant results are those most likely to be reported, distortion of information is very likely.

When these problems are considered, it is clear that some rigorous analysis of research results is needed before the findings can be incorporated in a research base.

Proposals

One step toward the preparation of a research base--which could serve the design and maintenance of teacher preparation programs as well as accreditation procedures--would be the establishing of an evaluation team for screening research relevant to teacher preparation. Such a team should include experts in research design and in teacher preparation as well as generalists who can take an overall view.

It seems unrealistic to expect each teacher preparation program to design, initiate, conduct, and analyze its own research studies, as Section 5 of the present *Recommended Standards* seems to require. Such studies would be seriously handicapped by a lack of the trained staff, the resources, and the mental set and orientation necessary for independent research. The duplication of effort would in any case be highly wasteful.

It is even doubtful whether any single organization--a teacher preparation program, or this Special Interest Group on the Teacher Preparation Curriculum, National Education Association, or Division B of the American Educational Research Association could marshall adequate resources to establish a research base for accreditation. By working in cooperation, these and other groups could develop a number of varied but rigorous and interconnected research designs, with procedures for processing the data. To such a collaborative effort, the teacher preparation programs could contribute the special knowledge and resources which they possess. The participation of researchers not connected with the programs would ensure careful design, rigorous procedures, and a minimum of bias. Replication could be ensured by making the same standardized research designs available to many programs. Variety and scope could be ensured by providing a number of different designs. Teacher preparation programs would benefit from accurate feedback about their graduates and from the economy with which they would obtain such information.

In this way we could build a solid research base for accreditation on replicated, carefully designed studies. Given the possibility of cooperative research, there seems to be no good reason why 1,100 teacher preparation programs in the United States, and hundreds more in the English-speaking world, should have to design their own research and follow-up programs. Every teacher preparation institution would, with relief, participate in a project which promised to help rather than to police its program.

Bibliography

1. Albrecht, G. "A Survey of Teacher Opinion in California," *Phi Delta Kappan*, 42:103-108; 1969.

2. American Association of Colleges for Teacher Education. "Research and Study in Teacher Education," *AACTE Bulletin*, 20; 1968.

3. ---. *Standards for Accreditation of Teacher Education*. (2nd draft). Washington, D.C.: the Association, 1968.

4. ---. *Recommended Standards for Teacher Education*. Washington, D.C.: the Association, 1969.

5. Beery, J. R. *Professional Preparation and Effectiveness of Beginning Teachers*. Coral Gables, Florida: University of Miami, Graphic Arts Press, 1960.

6. Bellack, A. A., and others. *The Language of the Classroom*. New York: Columbia University, Teachers College Press, 1966.

7. Beledsoe, F., and others. *Comparison Between Selected Characteristics and Performance of Provisionally and Professionally Certified Beginning Teachers in Georgia*. Final Report. Athens: University of Georgia, 1969.

8. Blosser, P. E., and R. W. Howe. "An Analysis of Research Related to the Education of Secondary School Science Teachers," *The Science Teacher*, 36:87-95; 1969.

9. Brase, P. C., Jr. *A Study of Curriculum Changes in Selected Two Year Colleges Initiated by the Evaluation Process Used by the Middle States Association*. New York: Columbia University, Teachers College, 1964.

10. Brown, B. B. *An Investigation of Observer-Judge Ratings of Teacher Competence*. Final Report. Gainesville: University of Florida, 1969.

11. Bruce, M. H. "Teacher Education in Science," *Review of Educational Research*, 39:415-27; 1969.

12. Butcher, H. J. "The Attitudes of Student Teachers to Education: A Comparison with the Attitudes of Experienced Teachers and a Study of Changes During Training Courses," *British Journal of Social and Clinical Psychology*, 4:17-24; February 1965.

13. California Teachers Association. *California Teachers' Reactions to Certification and Pre-Service Courses*. Supplementary Research Report No. 58. Burlingame: the Association, Department of Research, 1966.

14. Canadian Teachers Federation. *Pre-Service Teacher Education in Canada*. Bibliographies in Education No. 1. Ottawa: the Federation, 1969.

15. Charters, W. W., D. Waples, and S. P. Capen. *The Commonwealth Teacher-Training Study*. Chicago: University of Chicago Press, 1929.

16. Dussault, G. *A Theory of Supervision in Teacher Education*. New York: Columbia University, Teachers College Press [in press].

17. Eidell, T., and J. Klebe. *Annotated Bibliography on the Evaluation of Educational Programs*. Eugene: University of Oregon, 1968.

18. Evans, K. M. "Research on Teaching Ability," *Educational Research*, 1:22-36; 1959.

19. Evers, N. H. "By What Yardstick Shall Be Judged?" *Evaluative Criteria for Accrediting Teacher Education: A Source Book on Selected Issues*. Washington, D.C.: American Association of Colleges for Teacher Education, 1967. pp. 59-67.

20. Fishburn, C. E. "Learning the Role of the Teacher," *Journal of Teacher Education*, 17:329-31; 1966.

21. Fink, M. "Ethnocentrism as It Relates to Teaching Success," *California Journal of Educational Research*, 4:111-14; 1953.

22. Foshay, A. W. "Curriculum and Instruction," *In Dundee Team Teaching Project, Greenwich, Connecticut: A Research Report*. New York: Columbia University, Teachers College, Institute of Field Studies, 1965. pp. 71-176.

23. Frederiksen, N. "Proficiency Tests for Training Education," *Training Research and Education*, edited by R. Glaser. New York: Science Editions, 1965. pp. 323-46.

24. Heidelbach, R., and M. Lindsey. *Annotated Bibliography on Laboratory Experiences and Related Activities in the Professional Education of Teachers, July 1966-June 1967*. Washington, D.C.: Association for Student Teaching, 1968.

25. Herbert, J. "Direct Observation as a Research Technique," *Psychology in the Schools*, 1970 [in press].

26. ---, and D. P. Williams. "Psychology in the Teacher Preparation Program," *Psychology in Teacher Preparation*, edited by J. Herbert and D. P. Ausubel. Monograph Series No. 5. Toronto: Ontario Institute for Studies in Education, 1969.

27. Johnson, J. A. *A National Survey of Student Teaching Programs*. Final Report. DeKalb: Northern Illinois University; Washington, D.C.: Department of Health, Education and Welfare, Office of Education, Bureau of Research, 1968.

28. Johnson, M., Jr. "Profiles of Beginning Junior High School Teachers," *Journal of Teacher Education*, 16:303-06; 1965.

29. Johnson, W. D. "Microteaching: A Medium in Which To Study Teaching," *High School Journal*, 51:92-96; 1967.

30. Kersh, B. *Classroom Simulation: A New Dimension in Teacher Education*. Monmouth, Oregon: Oregon State System of Higher Education, Teaching Research Division, 1963.

31. Lindsey, M., and others. *Annotated Bibliography on the Professional Education of Teachers*. Washington, D.C.: Association for Student Teaching, 1969.

32. Lowery, L. F. "A Study of the Attitudes of Parents Toward Teachers," *Journal of Educational Research*, 62:227-30; 1969.

33. Lueck, W. R. "Professional Insecurity of Prospective Teachers," *Teachers College Journal*, 38:119-25; 1965.

34. MacDonald, J. B. "Researching Curriculum Output: The Use of a General Systems Theory To Identify Appropriate Curriculum Outputs and Research Hypotheses." Paper presented at the American Educational Research Association annual meeting, 1965.

35. McDonald, F. J., and D. W. Allen. *Training Effects of Feedback in Modeling Procedures on Teaching Performance*. Stanford, Calif.: Stanford University, Department of Education, 1967.

36. McGuire, C. H. "A Process Approach to the Construction and Analysis of Medical Examinations," *Journal of Medical Education*, 38:556-63; 1963.

37. ---, and D. Babbott. "Simulation Technique in the Measurement of Problem-Solving Skills," *Journal of Educational Measurement*, 4:1-10; 1967.

38. Maucker, J. W. "Imperatives for Excellence in Teacher Education," *Foundations for Excellence*. Fifteenth Yearbook. Washington, D.C.: American Association of Colleges for Teacher Education, 1962. pp. 1-12.

39. Metzner, S. "The Teacher Preparation Myth: A Phoenix Too Frequent," *Phi Delta Kappan*, 50:105-07; 1968.

40. National Education Association. "Teacher Mobility and Loss," *National Education Association Research Bulletin*, 46:118-26; 1968.

41. ---, Research Division. *Teacher Mobility and Turnover, 1965-66 to 1966-67*. Washington, D.C.: the Division, 1969.

42. National Union of Teachers. *The Future of Teacher Education*. London: National Union of Teachers, 1969.

43. Orlich, D. C. "Idaho Teacher Turnover: 1965--A Selected Analysis of the Problem," *Journal of Teacher Education*, 17:447-53; 1967.

44. Remners, H. H. "Rating Methods in Research on Teaching," *Handbook of Research on Teaching*, edited by N. L. Gage. Chicago: Rand McNally, 1963. pp. 329-78.

45. Rosenthal, R. "Empirical Versus Decreed Validation of Clocks and Tests," *American Educational Research Journal*, 6:689-91; 1969.

46. ---, and L. Jacobson. *Pygmalion in the Classroom*. New York: Holt, Rinehart and Winston, 1968.

47. Ryans, D. G. "Patterns of Teacher Classroom Behavior and Their Assessment," *Characteristics of Teachers: Their Description, Comparison, and Appraisal*. Washington, D.C.: American Council on Education, 1960.

48. Schalock, H. D., J. H. Beaird, and H. Simmons. *Motion Pictures as Test Stimuli: An Application of New Media to the Prediction of Complex Behavior*. Monmouth: Oregon State System of Higher Education, Teaching Research Division, 1964.

49. Simon, A., and E. Gil Boyer, eds. *Mirrors for Behavior: An Anthology of Classroom Observation Instruments*. Vols. 1-6. Philadelphia, Pa.: Research for Better Schools, 1968.

50. ---. *Mirrors for Behavior: An Anthology of Classroom Observation Instruments Continued*. Vols. 7-15. Philadelphia, Pa.: Research for Better Schools, 1970.

51. Smith, M. C. "The Case for Teachers Who Are Specifically Prepared To Teach in Junior High Schools," *Journal of Teacher Education*, 17:438-43; 1966.

52. Start, K. B. "A Follow-Up of the 1961 Group of Teachers," *Selection, Training and Professional Progress of Teachers*. Report No. 1. Manchester, England: Machester University, School of Education, 1967.

53. ---. "Teachers from Different Colleges of Education," *Selection, Training and Professional Progress of Teachers*. Report No. 2. Manchester, England: Manchester University, School of Education, 1967.

54. Steele, P. M. "Changes in Attitude Amongst Training College Students Towards Education in Junior Schools." Unpublished Master's thesis, University of Manchester Library, 1958.

55. Stern, G. "Measuring Non-cognitive Variables in Research on Teaching," *Handbook of Research on Teaching*, edited by N. L. Gage. Chicago: Rand McNally, 1963. pp. 398-447.

56. Stiles, L. J., and others. *Teacher Certification and Preparation in Massachusetts*. Boston: Massachusetts Advisory Council of Education, 1968.

57. Thorndike, R. L. "Review of *Pygmalion in the Classroom* by R. Rosenthal and L. Jacobson," *American Educational Research Journal*, 5:708-11; November 1968.

58. ---. "But You Have To Know How To Tell the Time," *American Educational Research Journal*, 6:692; 1969.

59. Turner, R. L. "Characteristics of Beginning Teachers: Their Differential Linkage with School System Types," *School Review*, 73:48-58; 1965.

60. Wandt, E. "A Comparison of the Attitudes of Contrasting Groups of Teachers," *Educational and Psychological Measurement*, 14:418-22; 1954.

61. Wiseman, S., and K. B. Start. "A Follow-Up of teachers Five Years After Completing Their Training," *British Journal of Educational Psychology*, 34:342-61; 1965.

62. Woodring, P. *New Directions in Teacher Education: An Interim Report of the Fund for the Advancement of Education in the Areas of Teacher Education and Recruitment*. New York: Fund for the Advancement of Education, 1957.

THE RESEARCH BASE FOR TEACHER PREPARATION CURRICULUM STANDARDS

S. C. T. Clarke
University of Alberta

The preparation of teachers includes many aspects, one of which is curriculum. The *Recommended Standards for Teacher Education* (1) (hereinafter called *Recommended Standards*) specify curricula; faculty; students; resources and facilities; and evaluation, program review, and planning. Engbretson (3), in this analysis of the proposals for model elementary teacher education programs which were not funded by the United States Office of Education, describes how the program components used in the request for proposals were developed initially by a team of consultants. The components were goals, selection, learning experiences, relationship of professional sequence to undergraduate program, inservice education, faculty, evaluation and feedback, management systems, and change mechanisms. The nine *Comprehensive Elementary Teacher Education Models* which were funded have been analyzed by a number of people, including Clarke (2), Shaftel (7), and Le Baron (4). The latter uses as headings for analysis: overview, instructional goals, overall program organization, curriculum design, institutional relationships, innovative features, student guidance, management and control, and placement and follow-up. It is clear that while differences of emphasis may occur, the main parts or aspects of teacher education are the ones used in the *Recommended Standards*. It is the differences in emphasis which constitute change in teacher education. Putting it another way, change is less likely to occur by the development and adoption of as yet some undreamed of innovation and is more likely to occur by altered emphasis among existing components.

This paper deals with curriculum in teacher education. The process of change just outlined is illustrated in the *Recommended Standards* being considered. In the 1967 draft, the first heading was "The Program of Instruction for Teacher Education." The 1969 *Recommended Standards* change this first heading to "Curricula for Basic Programs" and add two new sections: "Design of Curricula" and "Use of Guidelines Developed by National Learned Societies and Professional Associations." The extensive discussion, including written submissions, open hearings, and reports from discussion groups which preceded the 1967 version of the *Recommended Standards* and which followed its publication and served as a basis for the 1969 revision is described by Massanari (5).

The new section "Curricula for Basic Programs" contains these definitions:

> As used in the following standards, a "teacher education program" refers to the curriculum, the teaching, the learning, and the supporting resources for the teaching and learning process. "Curriculum" includes the courses, seminars, readings, laboratory and clinical experiences, and practicum as described under the general studies component and the professional studies component. A "program of study" refers to the sequence of courses, seminars, readings, laboratory and clinical experiences, and practicum selected for each student (1:3).

It will be useful to remind ourselves of the typical kind of outline used in thinking about school curriculum in order to compare its components with those proposed for teacher education curriculum.

Typically, school curriculum can be thought of in terms of the following:

1. Selection of curriculum, for example, needs of society, the needs of the individual, trends in technology, hence the objectives of the curriculum.
2. Organization, for example, sequence of order, logical, psychological, modules, units, etc.
3. Presentation, for example, didactic, discovery, inquiry, discussion, work units, etc.
4. Motivation, for example, need identification and fulfillment, developing interests, reinforcement devices, etc.
5. Evaluation, for example, formative and summative, procedures used, frequency or continous, etc.
6. Change, for example, feedback mechanisms, change agents, provision for formal review, etc.

This is only one way of thinking about school curricula. Basic components include objectives, content, organization and presentation, and evaluation.

An examination of the "Curricula for Basic Programs" reveals eight standards. To present the standards alone is to do some injustice to their spirit and content because each is amplified in paragraphs of explanation. The following paraphrases are intended to outline the curriculum for teacher education contained in the *Recommended Standards*.

1. Teacher education curricula are based on objectives reflecting the institution's conception of the teacher's role. The objectives cover five elements: general studies, content for the teaching specialty, humanistic and behavioral studies, teaching and learning theory with laboratory and clinical experience, and practicum.
2. There is a planned general studies component of at least one-third of the time, covering the symbolics of information, natural and behavioral sciences, and humanities.
3. The professional studies component included knowledge of the school curriculum to be taught plus additional background in this content.
4. The professional studies component includes the historical and philosophical bases (humanistic studies) and the psychological, anthropological, economic, and political science bases (behavioral studies) of education.
5. The professional studies component includes the systematic study of teaching and learning theory with laboratory and clinical experience.
6. The professional studies component includes a substantial practicum.
7. In developing curricula the institution gives due consideration to guidelines prepared by national learned societies and professional associations.
8. The design, approval, continuous evaluation, and development of teacher education are in the hands of a unit the majority of whose members are teacher educators.

The reader who expects to find in the *Recommended Standards* a blueprint for teacher education is doomed to disappointment. With about 1200 institutions preparing teachers in the United States, and with the current lack

of secure knowledge about the preparation of teachers, it would wrong, and indeed folly, to detail a prescription. Rather, the *Recommended Standards* outline a framework, or alternatively, describe the essential attributes.

The first standard, added between the 1967 and 1969 versions, is objectives. It will be recalled that the USOE call for proposals commenced with this item. School curricula commence with this. The importance of objectives is a current fad in education and is linked to evaluation, which starts from objectives. However, the important point about the objectives for teacher education as called for in the *Recommended Standards* is that they must be derived from the institution's conception of the teacher's role. B. O. Smith (8) describes the derivation of teacher education curricula from a task analysis of teaching in these steps.

1. Analyze the job of teaching into the tasks that must be performed.
2. Specify the abilities required for the performance of these tasks.
3. Clearly describe the skills or techniques through which the abilities are expressed.
4. Work out in detail training situations and exercises for the development of each skill.
5. Classify training situations and exercises by tasks, abilities, skills, grade level, field of instruction, and backgrounds of children.
6. Establish training complexes.

If there is any weakness in the way the *Recommended Standards* deals with objectives, it is that of undue compression. The emphasis to be given to task analysis of teaching, and objectives derived from this, does not come through clearly in the *Recommended Standards*.

Returning to the standards listed for the "Curricula for Basic Programs" one notices an emphasis on knowledge: general studies, subject to be taught, humanistic and behavioral foundations. This has been the traditional emphasis. This emphasis is currently being seriously challenged by more and early field experiences, by micro-teaching and simulation in college classes, and by attempts to develop performance criteria and competency based curricula. The order in which the standards appear might cause the reader to conclude that knowledge is to receive more emphasis than the "systematic study of teaching and learning theory with laboratory and clinical experience, and a substantial practicum." A point previously made is here reiterated: The *Recommended Standards* include the essential attributes of teacher education. Change will likely be in terms of emphasis. The trend today is to emphasize performance.

THE RESEARCH BASE FOR TEACHER EDUCATION CURRICULA

The discussion so far leads us to the possibilities for a research base. An analysis of teaching soon reveals that teaching has many forms or facets: Explaining is different from leading a group discussion and from stimulating enquiry. The attempt to analyze teaching will seek to isolate the variables involved, and so will lead to discovery of commonalities or general features, and unique features. When the variables are isolated and their relationships stated, teaching theories, with perhaps one general theory and a number of specific theories, will develop. Side by side with the theory development (probably preceding it) will be studies based on small manageable units such as microteaching and simulation. Studies based on the observation of classroom interaction will no doubt continue. The sequence: objectives (in terms

of rather specific behaviors) then learning experiences (in terms such as microteaching and simulation) then analysis of process then evaluation of results then feedback, constitutes the theory then practice then evaluation then feedback loop which will not only improve the curriculum of teacher education but which will provide the research base for it.

There is evidence that the American Association of Colleges for Teacher Education has access to the latest theory and research in teacher education during the preparation of the *Recommended Standards*. Mars (6) describes its activities for dissemination and adoption of recent research in teacher education. These included the following topics and individuals: Analysis of Teaching, Dwight W. Allen and Ned A. Flanders; Nonverbal Communication, Charles M. Galloway; Simulation, Donald R. Cruickshank; Structural and Learning Analysis, Asahel D. Woodruff; and Media Analysis, Vernon S. Gerlock. A number of regional workshops were held with the avowed purpose of disseminating and hopefully securing adoption of the latest theory-based and research-based practices. Such activities no doubt influenced not only the staff of AACTE, but also the many educators throughout the country who had a voice in determining the final version of the *Recommended Standards*.

To return to a point previously made: With the present state of secure knowledge about teacher education, and considering that there are some 1200 institutions in the United States which prepare teachers, the *Recommended Standards* must necessarily provide a framework and not a blueprint. This is not to say that the framework can ignore the latest theory and research in teacher education. There is evidence to sho that the *Recommended Standards* reflect what was known in the late 1960's. However, the pace of theory development and research in teacher education is so great in the present state of ferment, that there have been many changes of emphasis. The only way to keep the *Recommended Standards* up to date is to have planned, periodic (say every five years), and thorough revision.

Bibliography

1. American Association of Colleges for Teacher Education. *Recommended Standards for Teacher Education*. Washington, D.C.: the Association, 1969.

2. Clarke, S. C. T. "The Story of Elementary Teacher Education Models," *Journal of Teacher Education*, 20:283-93; 1969.

3. Engbretson, W. E. *Analysis and Evaluation of Plans for Comprehensive Elementary Teacher Education Models*. Final Report. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education, Bureau of Research, 1969.

4. Le Baron W. *Analytical Summaries of Specifications for Model Teacher Education Programs*. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education, Bureau of Research, 1969.

5. Massanari, K. "The AACTE-NCATE Feasibility Project: A Test of Proposed New Accreditation Standards for Teacher Education," *Journal of Teacher Education*, 20:5-13; 1969.

6. Mars, Walter J. *Professional Teacher Education: A Program Design Developed by the AACTE Teacher Education and Media Project*. Washington, D.C.: American Association of Colleges for Teacher Education, 1968.

7. Shaftel, Fannie R. *The Stanford Evaluation of the Nine Elementary Teacher Training Models*. Final Report. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education, Bureau of Research, 1969.

8. Smith, B. O. *Teachers for the Real World*. Washington, D.C.: American Association of Colleges for Teacher Education, 1969.

THE ACCREDITATION RESEARCH BASE RELATIVE TO FACILITIES

Donald M. Medley
University of Virginia

The following statement appears in the introduction to the *Recommended Standards for Teacher Education*.

> NCATE accreditation validates the quality of preparation programs and signifies that persons recommended by the institution can be expected to perform satisfactorily in typical teaching and other professional school positions throughout the United States (2:1).

This statement clearly implies that if a program meets the recommended standards and is therefore accredited by the National Council for Accreditation of Teacher Education (NCATE), teachers trained in it will be competent--or, at least, will be more competent, on the average, than teachers trained in programs which do not meet recommended standards and are therefore not accreditated.

If the standards are to be based on research findings, then the research must show a correlation between certain program characteristics and measures of teaching competence or success of graduates of the program.

My discussion is on the research base underlying the *Recommended Standards* related to resources and facilities. This statement seems to me to define pretty clearly the kind of research base needed. Research evidence is needed that graduates of accredited programs--of programs which measure up to the recommended standards, that is--make better teachers than graduates of programs which do not measure up to these standards. The inclusion of any given standard should be justified by evidence that that particular element is related to the effectiveness of teachers graduated from the program which contains it.

I take this to mean that I should look for the following kinds of research: (a) studies showing that the quality of teachers is related to the quality of the library in the undergraduate colleges in which they receive their basic teacher training; (b) evidence that teachers graduating from colleges with adequate faculty offices or well-lighted classrooms tend to succeed better than teachers from colleges where classrooms and offices are antiquated or crowded--even though curricula, quality of faculty, and other characteristics of their training do not differ. I should review studies relating how much pupils learn from a teacher to what kinds of media were available in his undergraduate college.

Studies such as these are hard to find. If the validity of the standards for resources has to be based on research of this type, there is no basis for calling them valid. The fact that they do have validity is a result of the wisdom of the committee who drew them up--of their experience, their intuitive judgment, and their familiarity with what little research there is. And I am afraid that among these three, research makes by far for the smallest contribution.

Rather than picking over what bits and pieces there are, I shall discuss what might be done to close up the gap between needed research and what is available. Our AERA special interest group commitment is to the idea that decisions made in education in general, and in teacher education in particular, should be based on scientific knowledge, and that the members of committees like the one which drew up the *Recommended Standards* should be able to base their recommendations to a much greater extent on research than on either experience or intuition.

THREE WAYS TO ATTACK THE RESEARCH PROBLEM

As I see it, we have available to us three ways of attacking the problem of bringing research results closer to decision needs in teacher education. One thrust would employ careful, well-designed research; a second thrust would involve the systematic exchange of program evaluation data; and the third would take the form of a one-time crash project which I will describe later.

Thrust 1: Research

We would continue to encourage carefully designed experimental studies, most of which would be moderate in size and modest in purpose, in the sense that any single investigation would be directed only at a portion of the problem. The attack would be piecemeal and the increase in knowledge gradual.

The notion that program facilities contribute to competence of program graduates, for example, depends on a series of relationships something like this: (a) The presence of facilities is related to the nature of the experiences a teacher has in training. (b) The nature of the training experiences a teacher has is related to the terminal behaviors or skills he possesses when he graduates. (c) The terminal behaviors a graduate possesses are related to his success in the classroom--that is, his competence on the job.

The kind of research I have in mind would normally study relationships between contiguous elements in this chain. Some studies would attempt to establish relationships between institutional resources and curriculum innovations. Others would relate curriculum innovations to terminal behaviors. Still others would relate terminal behaviors to effectiveness in the classroom. Over a period of time, as more and more knowledge about such relationships accumulated, the chain of relationships from facilities to teacher competence would become stronger and stronger.

But the process would be a slow one, with relatively little pay-off in its early stages. Substantial pay-off would not come until a critical point is reached--the point at which it is possible to propound a theory broad enough to cover the entire series. Until this is done, bits of knowledge must be added one by one. Afterward, each increment of knowledge would reverberate throughout the system and its effects would be correspondingly multiplied.

As far as teacher education goes, it will be some time before we reach such a critical point. This does not mean we should slow up our research effort. On the contrary, we must continue to support all the research we can. We cannot really begin to make progress toward satisfactory solution to any of our problems until we have at least one viable research-based theory of teacher education.

Thrust 2: Information Exchange System

We would make more systematic use of the experience we gain in ongoing program operation. One of the things I like best about the new *Recommended Standards* is their emphasis on evaluation, program review, and long-term planning. A program which meets the standards recommended in this area has within itself a mechanism for continuous growth and self-improvement provided by a system of feedback. I propose to add a provision for cross-fertilization between programs through a systematic exchange of this kind of information. This would be operated through a central data bank with a special system for information storage and retrieval. There are doubtless a number of such systems which might serve, of which I shall suggest only one for purposes of discussion.

I would suggest that a special instrument be constructed on which the entire experience of one individual teacher education student may be recorded in a form which is easily accessible to a computer. Each student's record should include (a) data obtained at the time of admission, such as test scores and biographical data; (b) data regarding his experiences during training--possibly in the form of courses taken and credits earned; (c) all available evaluation data obtained during training--course grades, test scores, ratings by supervisors, etc.; and (d) whatever follow-up data are obtained by the institution after the student graduates, especially those related to success or failure as a teacher.

I envision such data as being precoded and entered by the institution on special mark-sensing forms from which it may read onto master tapes by a Digitek, IBM 1230, or the like, and stored in a central data bank.

Once the bank has been set up, a user could test hypotheses about treatments by asking the bank questions of the form: "How have teachers with characteristics A, B, etc., who have had experiences X, Y, etc., performed (a) in training, and (b) on the job?" The inquirer could specify both the type of teacher and the nature of the experiences in as much detail as he chose. The computer would search all the records from all of the participating institutions and produce a report summarizing all relevant information about such teachers and, if desired, about a comparable group who had not had the experience as well.

The success of the proposed scheme would depend on the development of a recording procedure which would not place an undue clerical burden on the institution and yet would yield accurate and usable data. We have developed, in connection with a large-scale longitudinal study of early childhood education, an analogous information system. We have been recording the in-school experience of school children rather than college

students. The child's experience is recorded by a non-professional recorder on a machine-scorable answer sheet from which it is read onto magnetic tape and collated with other records of the same child. The complete record constitutes something very like a narrative account of what has happened to the child during his first five years in school. There is a program which enables us to test hypotheses in the data in much the way we have proposed for the teacher education system, by merely punching the question on a parameter card and running it through the computer. We have considerable flexibility both as to the type of question we may ask and the kinds of records to be used in answering it.

I am in no sense of the term a systems expert, but experiences seem to indicate that this kind of operation is technically feasible. Whether or not it would be possible to get any sizable number of institutions to participate in such an operation is another matter. The potential value of the information available to each participating institution would seem to me to justify the trouble and expense involved. If the operation were managed by an agency such as the American Association of Colleges for Teacher Education, it could also provide information highly relevant to such activities as the setting of standards for accreditation, at little extra expense.

Thrust 3: A Crash Program

Neither of the two strategies I have described can be expected to yield results that would be immediately useful as a basis for policy decisions such as setting accreditation standards. To provide something to use while we are waiting I am going to suggest a crash program designed to exploit the information available in current practice in teacher education--a one-shot affair (which some might describe as quick and dirty). I am talking about a large-scale questionnaire study patterned to some extent after the Equal Opportunities Survey, but focused on teacher education institutions rather than on the schools.

Two kinds of data would be collected from two different sources. From every institution in the country that trains an appreciable number of teachers, data would be collected about program characteristics of all kinds. Simultaneously, from first- and second-year teachers in a representative sample of the nation's schools, data would be collected related to their success in teaching together with descriptive information about their preservice training--where it had been obtained, what specific courses were taken, and so on.

Modern statistical procedures are available which would make it possible to extract from these data estimates of relationships between specific elements in the programs and the success of graduates of those programs, partialing out such things as the school situation in which a teacher is placed, variations in course patterns pursued by different teachers who are trained in the same basic program, and program characteristics other than the one being studied.

Despite these controls it is likely that relationships would be severly attenuated by uncontrolled sources of variation. The study

would therefore need to be sufficiently large in size so that even moderate relationships could be detected reliably. Caution would need to be exercised in inferring casual relationships, and the data could scarcely be regarded as of high quality.

However, the findings of such a study may be expected to shed light into many dark corners. Besides fulfilling their main purpose of providing direct evidence as to the validity of the proposed standards, the data should be rich in leads for research studies tc be done under Thrust 1 and for program innovations to be tried under Thrust 2. I can conceive of no other way of moving so far so fast toward solving so many perplexing problems.

Some Conclusions

Ever since I can remember I have heard research workers criticizing practitioners for not basing their decisions on research, and practitioners criticizing research workers for not developing useful and unambiguous information on which decisions could safely be based. Although I have not actually heard any recrimination from either group in relationship to the *Recommended Standards*, they do provide an excellent, and a disturbing example of such a gap--particularly in regrad to my own assignments: standards of resources and facilities.

One can scarcely blame the committee which drew up the standards for using all the wisdom and experience they could command in doing so. Indeed, one can only admire them for the thoroughly workmanlike job they have done and the valuable tool they have forged for improving teacher education. If a researcher looking at them has any reaction,it tends to be one of rueful admiration at how well they seem to do in an area where research is of so little assistance.

Nor do I think it fair to blame us research workers for not having provided more data relevant to the specifc problems the committee faced, since we never offered to do the job in the first place. We have been too busy with a quite different problem: That of building a scientific basis for teacher education. As so often happens, we have attended to the important at the expense of the merely urgent.

What I have tried to do today is to suggest that we clearly discriminate between two kinds of activity, both of which might be referred to as research in teacher education, since both are designed to improve it, but which differ sharply in their intermediate objectives--that is, in the strategies or means they employ to achieve their common ultimate goal.

I have suggested that we make a clear distinction between the two activities not so that we may press the one and neglect the other, but so that we may support and encourage both of them, and know when we are doing which and what to expect of each. I have also suggested how we might deploy our resources most effectively between them in the immediate future.

Bibliography

1. Allen, D., and K. Ryan. *Microteaching*. Reading, Mass.: Addison-Wesley, 1969.

2. American Association of Colleges for Teacher Education. *Recommended Starndards for Teacher Education*. Washington, D.C.: the Association, 1969.

3. Barr, A. S., and A. Singer. "Evaluative Studies of Teacher Education," *Journal of Teacher Education*, 4:65-72; 1953.

4. Benton, C. W., and others. *Television in Urban Education*. New York: Frederick A. Praeger, 1969.

5. Bosley, H. E., ed. *Teacher Education in Transition*. Vol. 2. Baltimore, Md.: Multi-State Education Project, 1969.

6. Bowers, H. *Research in the Training of Teachers*. Toronto, Canada: J. M. Dent and Sons and MacMillan, 1952.

7. Bradfield, L. E. "A Survey of Twenty-four Campus Elementary Schools," *Journal of Teacher Education*, 6:118-21; 1953.

8. Childs, J., and others. "The Uses of Television in Teacher Education," *Audio Visual Instruction*, 12:1050; 1967.

9. Cyphert, Frederick R., and Ernest Spaights. *An Analysis and Projection of Research in Teacher Education*. Columbus: Ohio State University Research Foundation, 1964.

10. Ellis, Eleanor V. *The Role of the Curriculum Laboratory in the Preparation of Quality Teachers*. Tallahassee: Florida A&M University Foundation, 1969.

11. Haskew, L. D. "Mechanical Controls of Quality in Teacher Education," *Journal of Education*, 1:121-27; 1950.

12. Heidelbach, Ruth, and others. *Annotated Bibliography on Laboratory Experiences and Related Activities in the Professional Education of Teachers*. Washington, D.C.: Association for Student Teaching, a national affiliate of the National Education Association, 1967.

13. Isler, N. P. *Planning in Higher Education: An Interpretive Bibliography*. Madison, Wis.: ERIC Clearinghouse on Educational Facilities, 1969.

14. Kallenbach, W. W., and M. D. Gall. "The Comparative Effects of Microteaching and a Conventional Training Program in the Classroom Performance of Elementary Intern Teachers," *Journal of Educational Research* [in press].

15. Koerner, J. D. "Findings and Prejudices in Teacher Education," *School and Society*, 92:127-34; 1964.

16. LaGrone, H. F. *A Proposal for the Revision of a Pre-Service Professional Component of a Program of Teacher Education*. Washington, D.C.: American Association of Colleges for Teacher Education, 1964.

17. Lindsey, Margaret. *Annotated Bibliography on the Professional Education of Teachers*. Washington, D.C.: Association of Student Teaching, a national affiliate of the National Education Association, 1969.

18. Newton, D. E., and F. G. Watson. *The Research on Science Education Survey*. Cambridge, Mass.: Harvard University, 1968.

19. Pierce, Alice A. "Individualizing Teacher Education," *Journal of Teacher Education*, 4:118-21; 1953.

20. Rogers, W. R. *Television Utilization in the Observation Programs for Teacher Education*. San Jose, Calif.: San Jose State College, 1963.

21. Schueler, H., M. J. Gold, and H. E. Mitzel. *The Use of Television for Improving Teacher Training and for Improving Measures of Teacher Performance*. New York: City University of New York, Hunter College, 1962.

22. Schueler, H., and G. Lesser. *Teacher Education and the New Media*. Washington, D.C.: American Association of Colleges for Teacher Education, 1967.

23. Shuster, A. H. "Professional Education and Teaching," *Journal of Teacher Education*, 6:258-62; 1955.

24. Smith, Lillian C. "A Study of the Use of Micro-Teaching in the Preparation of Elementary Teachers," *Teacher Education in Transition*, Vol. 2, edited by H. E. Bosley. Baltimore, Md.: Multi-State Teacher Education Project, 1969. pp. 121-35.

25. Smith, L. M., and P. M. Keith. "Fantasy and Reality in the Language of the New Technology," *Educational Technology*, 5-9; 1968.

26. ---. *Social Psychological Aspects of School Design*. Final Report. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education, Bureau of Research, 1967.

27. Umruh, Adolf. "What's Needed in Teacher Education?" *Phi Delta Kappan*, 37:258-61; 1956.

28. Webb, C. D., and others. "Progress Report on an Experiment Reducing In-Class Student Teaching Time and Substituting Micro-Teaching Sessions," *Proceedings of the Utah Academy of Arts, Sciences, and Letters*, 44:224-28; 1967.

STANDARDS OF STUDENT SELECTION AND INTERACTION

John R. Dettre
University of Kentucky

This paper will attempt to focus on two things: (a) the current status of the literature related to the standards for admissions, selective retention, and student involvement in program development and evaluation; (b) identified areas in need of additional research in relation to these same three areas in the standards. Few, if any, studies really exist that involve the substance of the standards directly. What is presented, therefore, is relative and reflects an inferential and judgmental process on the part of the writer.

Since the paper proceeds from a inferential and judgmental base, it will be necessary for the reader to make his own decisions and draw his own conclusions using his own frame of reference. The reader should consider a number of factors in attempting to gain some perspective on the treatment of information in this portion of the standards.

GAINING A PERSPECTIVE

One factor to be considered is the matter of standards. Stated simply, a standard is anything used as a basis for comparison. It is assumed that standards are developed to permit one to make comparisons between or among individuals or groups, or to compare individuals or groups with some predetermined and minimally desirable level of excellence. In the case of accreditation, it is assumed that standards are developed for the purpose of determining the extent to which a given institution fails to meet, equals, or exceeds some minimally desirable level of excellence in selected areas.

Standards, however, may be provided in three or more ways: (a) They may reflect a series of activities in which an institution is expected to engage. (b) They may indicate various quantities or degrees of given states of affairs one should possess or provide. (c) They may indicate both a series of activities and minimal quantities or degrees an institution is expected to produce in relation to those activities. Standards related to admissions, selective retention, and student involvement in program development and evaluation reflect an indication of a series of activities in which an institution is expected to engage.

A second consideration may be illustrated best by example. If one is interested in selecting a parachute and intends to use the findings of research on different brands of parachutes in selecting one, then it is important that the selection be supported by research findings reflecting the highest level of significance possible. If one is afflicted with cancer and in search of some cure, then any level of significance for almost any possible treatment may be acceptable. In considering the literature related to this portion of the standards, the question of level of significance of the findings one will accept will be an individual matter and will reflect the amount of urgency and need the reader perceives in relation to the standards.

The matter of priorities constitutes a third consideration. It is no doubt more comforting and convenient to believe one can and does deal with the matter of standards by investigating all parts concurrently. The fact remains, however, that some parts of the standards are given positions of greater and less importance. It is important to recognize and understand the basis used to make these distinctions. In particular it is important to reflect on the concept of critical functions and priorities as he responds to the whole matter of admissions, selective retention, and student interaction in the program.

Whatever the reasons given, a fourth consideration is that a high priority has been given to the identification of teacher competence or effectiveness. Such identification leads to improved ability on the part of institutions to select students and educate them in the ways of teaching. Part 4 response to current investigations of the standards will undoubtedly reflect the degree of agreement or disagreement he has with the priority given this identification process. Even for those who agree with the priority there is still another consideration to be made: The extent to which the teacher effectiveness is to be determined in relation to ultimate rather than proximate learning outcomes (39; 52).

A fifth consideration involves the extent to which the standards should reflect a preference for individual as opposed to institutional or professional expectations. Individual expectations require standards based on maximizing the uniqueness of each individual. Institutional or professional expectations require the identification and development of both common experiences for all candidates and the use of assessment practices based on comparisons among candidates. The implications are extremely important for the standards and program development, depending upon the position taken.

Finally, there is the personal matter of definitions, substantively and operationally, to use in treating the literatures. For example, at one point in time, the efficacy of recruitment procedures as a part of a working definition of admissions could be found. Recruitment now seems to be a matter of history. There is also the matter of how inclusive an individual wants to be in defining many of the terms in the standards. Some may wish to use definitions they feel are quite literal. Others may wish to combine or otherwise blend two or more activities into one operational definition. The differences in definition will also make a difference in response to this treatment of the standards.

In summary, there is a need to develop some notion of a position regarding the scope, sequence, and substance of those activities which are significant parts of the section under consideration here. The six considerations suggested above are important to the formation of such a position.

DEFINITION OF TERMS

To facilitate communications and to provide some baseline, the following definitions of terms served are provided. The definitions were used in organizing the subsequent pages.

Admissions, as used here, refers only to the initial selection process. Included in the process are the establishment of selection criteria, collection of data, and the decision making process employed to select some students for the program and to reject others.

Selective retention, as used here, refers to a process of continuous evaluation of those already admitted to the program for the purpose of further refining the quality of teaching candidates. Included is the study of the individual's development of stated skills and competencies that constitute a continuum of learning outcomes.

Student involvement in the program development and evaluation, as used here, refers to a deliberate and planned program of interaction between students and the institution for analyzing and assessing existing programs, and for developing new programs. Included would be student responses, oral and written, to the nature and extent of the objectives, content, and methodology in existing programs, and the inclusion of student members on those committees and study groups dealing with curriculum and instruction.

Student teacher, as used here, refers to the student seeking to gain certification as a teacher. The term is not restricted by definition to those engaged solely in "student teaching" in the laboratory phase.

LIMITATIONS

The substance of this paper is essentially a study of studies. As a result, findings and conclusions are not based on an exhaustive examination of all available literature. Rather, those studies included are illustrative.

The paper is further limited by both the quality and inferential potential of both the studies used and the standards themselves.

Finally, those areas suggested as areas in need of additional research largely reflect others writing in the field. No claim is made either for the urgency or presumed utility of studies proposed.

ORGANIZATION

The remainder of this paper will be divided into two broad presentations. The first part will include a description of findings judged to be illustrative of the current level of information. The second part will include recommendations for further study.

THE STANDARDS

This portion will describe what the literature seems to report regarding admissions, selective retention, and student involvement in program development and evaluation. Few if any studies exist that treat the substance of the standards directly. The nature of the following, therefore, is relative and depends on inferential and judgmental processes.

Admissions

Admissions have been divided into four components: (a) criteria, (b) data collection procedures, (c) decision making activities in admissions, and (d) general.

Criteria. In order to operate an admissions program, it must have some criteria to use in drawing comparisons either among students, or between students and some minimally desirable model of a prospective applicant. Edson (30), McGee (47), O'Donnell (58), Stout (75), among others, indicate the use of grade point hour, selected course grades, and some measured competency in oral and written expression in most criteria for admissions. O'Donnell (58) also found the emotional stability is identified in some programs as a criterion. Lucio and others (46) have pointed toward the inclusion of more psychophysiological factors in admissions criteria. While Lucio is concerned with developing predictive indices to help identify people who ought to be directed toward some other occupational goals, it follows that if the use of psychophysiological factors will permit the kinds of distinctions suggested, then the same factors should become effective items in admissions work. In addition, there are isolated instances where one may also find measures of attitudes or belief systems as suggested criteria in admissions work (58; 70).

While work using grades or grade point averages, for example, may be useful as internal checks on the kinds of students one has in a program (31; 32), the evidence suggests that grades, as well as personality or attitudinal measures, still have no predictive validity (20; 24; 60). Shaver (70) did find a correlation between open-closed mindedness and selected teacher characteristics, but he conceded his findings were of little value unless the relationships could be confirmed in classroom situations. There is little to suggest that there is anything different now regarding grades as predictors from the conclusions drawn by Domas and Tiedeman (27) or conclusions reached from similar reviews of literature as, for example, in the case of Mascho and others (48). One inevitably seems to come back to the decades-old problem of general agreement on the idea of admissions but lack of significant and consistent evidence to support any criteria for an admissions process. Mascho and others seem to sum up the current state of affairs nicely where current and typical criteria for admissions are used:

> Despite decades of research on the problem, educators must face the fact that there is no common agreement on describing or evaluating teacher competence. Further, it is one thing to assert that a teacher should possess cheerful, friendly and sympathetic characteristics rather than their opposites, but quite another to identify in objective terms the specific and distinctive qualities of an effective teacher (48).

In addition to this problem, there is the whole unresolved matter of ultimate versus proximate behavioral outcomes. This must be settled if criteria are ever to be established (39).

There are other problems of note quite apart from being able to describe the characteristics definitively. McClure (51) reminds us that while personality may be an acceptable label for some potential selection criteria, the eventual relationship--even if it is established --will have to mediate objections likely to follow from those both inside and outside the profession.

Cook examined another dimension often ignored in research when he studied personal data and relationship to success in teacher education programs and entry into teaching. As Cook reasons, if such data are collected, then they must serve some valuable purposes. Although the study is limited in terms of its inferential potential, Cook's findings raise serious questions as to the real significance of personal data for matters of admission or retention.

Part of the problem seems to be centered around the validity and reliability of grades. As some of the studies suggest, there appear to be efforts to circumvent the use of grades in the belief that they are too faulty themselves to provide any useful base for projections or viewing predictive validity (46; 70). Perhaps the direction taken by Muro (55) points to one of the most important considerations. In his study the key variable seemed not to be the grade earned but the amount of convergence between the thrust of a course and the concerns of students in the course.

The notion of concerns opens other avenues for possible investigation, not only in relation to grades but in the whole relation between concerns and the ability of students to benefit from a preparatory program. In particular, the work of Fuller (34; 35) holds real promise for further exploration. Certainly the basic rationale for responding to concerns, and the conceptualization provided, gives promise of producing fruitful evidence on both admissions and selective retention.

If some of the findings involving different modes of counseling can be supported by future efforts (9; 38; 59; 64; 74), or continued results such as those obtained by Koran (43), Hart (36) and McCall (50) in matters of modifying actual classroom behavior can be projected, then it may be possible to discontinue institutional or professional programs of admissions and make a valid assumption that it is possible to prepare all to be the kinds of teachers they are capable of becoming. And should individualization prevail, then the notion of selection ratios (65) will assume greater significance in admissions work.

There is some indication, as for example in the case of Ort's effort (60) and Cohen's undertaking (22), to suggest that the best predicators of future success have been descriptive analyses of student teachers provided by cooperating teachers and university coordinators. More than three decades ago, Beeley (8) presented the idea of clinical techniques in selecting prospective teachers. The evidence gained from laboratory experiences, combined with the counseling and individual learning experiences identified earlier, might suggest much greater attention to a form of performance screening as the basic technique in admissions work.

In summary, no real research exists to test the validity and reliability of admissions programs. The evidence does suggest the existence of diversified programs with some commonly used criteria like grade point average, grades in particular courses, and competency in both the oral and written aspects of English. Further, there is little if any evidence to suggest that any or all such criteria do permit institutions to predict future success above the level of chance. There is still the unresolved notion of whether measures of teacher competency are to be derived from proven relationships to ultimate rather than proximate kinds of learning outcomes. Finally, such matters as the role of student concerns, the concept of performance screening as the basic tool, and the significance of individual over institutional expectations, remain a state of relative uncertainty in matters of admissions and criteria for admissions.

Collecting admissions information. While no actual count is provided, one may infer that most institutions: (a) have students complete a personal data sheet, (b) maintain a cumulative record of grades and test scores, (c) conduct something loosely called an interview as their basic procedures for collecting information about applicants. There is isolated evidence--the University of Pittsburgh (81), Florida State University (33), University of Georgia (79), the University of New Mexico, and the University of Kentucky--to suggest something of a beginning of a more complete data bank on those seeking admissions. None of these efforts, however, have been combined or expanded into more concerted efforts to date under the auspices of some central agency. Until they are, there is very little prospect for them to provide significant help to the membership-at-large.

The literature suggests that procedures collecting information are based on conventional wisdom in most institutions. Institutions tend to resort to methods presenting the greatest ease and convenience for them.

Perhaps the most discouraging thing in viewing efforts in data collection to date is the limited number of longitudinal studies completed or reported as being in progress. The efforts of Lucio (46) or Mascho and others (48) provides useful and practical examples for those more interested in exploring procedures for data collection than in building expansive bibliographies.

Decision making in admissions. Simply collecting data in relation to established criteria is not all there is to admissions. To do an effective job in admissions, the decision making process employed must also make effective and efficient use of the information collected. In those cases where minimally desirable levels of excellence are easily quantitied (31), the matter of the decision to be made would seem inconsequential. If, however, the information provided is both continuous and dichotomous and the criteria incorporate an expanded number of variables, then how the institution makes admissions decisions becomes a major consideration.

Bolton (11) examined the variables involved in decision making in the selection of teachers for teaching jobs. Although his work is not

directly related to standards and admissions, it is judged to be relevant. As his study viewed the process, there are four dependent variables in decision making and selecting personnel: (a) time, (b) discrimination, (c) certainty, and (d) consistency. Bolton's remark, "Unless decisions are both discriminative and consistent, there is little foundation upon which to accumulate evidence as to the validity of the decisions being made," sums up the importance of the decision making process in programs of teacher education (11). Bolton found that what kinds of information you use makes a difference in terms of the four dependent variables identified. He also provides a useful system of categorizing information by suggesting documents, interviews, and masked data (i.e. statements of exceptionalities) as possible labels.

There seems to be a number of factors that impinge upon the decision making process. Time and finance are two of the more obvious. Internally, however, the decisions must reflect valid and reliable distinctions among those selected and those rejected. The distinctions must be consistent. All those interested in the future of education must be assured that the decisions made are actually selecting people who will" . . . have the ability and interest to carry a college program to completion and the personal characteristics to make a successful career in school work (29)."

As a legitimate part of the triad of admissions work--establishing criteria, collecting meaningful information, and making decisions for selection and rejection--the area of decision making seems quite neglected. As the efforts to select candidates become more complex, the relative importance of decision making demands that more study be devoted to the nature of the process in admissions work in teacher education programs.

General. Before considering selective retention, there are one or two general matters that need to be identified. One is presented by Rabinovitz and Mitzel (63). It involves the notion of a selection of a ratio and its impact on admissions work. As the authors suggest, a selection ratio is not an arbitrary value; it cannot be established without considering the relative size of " . . . the applicant and to-be-selected groups (63)." When numbers exceed openings, the ratio is low. When they are equal, or openings exceed numbers, the ratio is high. As the authors demonstrate, when schools were begging for teachers, making distinct and discriminate decisions in admissions was a myth. The only concern was getting enough people out to meet the demands. The selection ratio was high. If the educated guesses on manpower needs for the 1970's are accurate, the selection ratio will be lowering. The implications for admissions are important. There is little evidence to suggest, however, that selection ratios have been seriously treated in relation to standards on admissions.

The other general matter is the development of complete admissions programs capable of taking candidates from their initial contact through to the actual placement in the first phases of a program. The University of Pittsburgh (81) has one possible model that seems worthy of consideration in its Comprehensive Elementary Teacher Education Model.

The need for a complete admissions package should not be ignored as a search is made for criteria, data collection procedures, and decision making processes.

Selective Retention

Selective retention has been divided into (a) student self selection, (b) instructional influence, (c) program influence, and (d) behavior modification.

Student self selection. In an ideal sense, students might be able to objectively analyze themselves and either remain in a program or drop out. There is little to suggest that students actually do this. One may infer, however, that other forces and factors quite removed from the teacher education program do lead students already admitted to drop out before they complete a program. Notestine (57) found those who withdraw differ from those who continue in three personal areas: (a) lonesome and unhappy, (b) discouraged by low grades, and (c) lack of interest in studies. He also found that these same people had the lowest high school class rank for their group, and lowest Scholastic Aptitude Test scores in both the verbal and math areas. It seems more likely, however, that developments along the lines pursued by Lucio and others (46), as well as counseling efforts (59), may yield a more productive return in deciding who should and who should not continue in a program of teacher education. Left to their own devices, and assuming grades can be maintained, conventional wisdom suggests few if any students will select themselves out of a program.

Instructional influence. The overall report provided by Bush and Gage (17) of developments and conclusions reached in the center at Stanford strongly supports both the notions of instructional and program influence in matters of selective retention. One can infer from the description that both areas provide extensive resources of valuable data. Furthermore, both aspects, instructional techniques and program organization, created a necessary baseline of defensible information to allow students and others to get a more objective and realistic perception of themselves in teaching. For example, the approach studied by Koran (43) pointed to the aptitudes of the students as they relate to the acquisition of a teaching skill, and students involved could see clearly just how they were developing in relation to the skill being taught.

Using a somewhat different approach, Hart (36) found that by changing the kind of instructional techniques used from conventional lecture-discussion to the use of discussant-stimulants he could produce significant gains in attitude. In fact, the very nature of the approach used suggests, again, that through more direct involvement of students in their own learning, more objective and defensible profiles of students as prospective teachers can be obtained. Such profiles in turn should permit the student to make a better decision on his own as to whether or not he should continue in a program. McCall (50) in another approach using dissonance and coded feedback was able to modify behavior and influence the self-perceptions of students.

The extent of simulation in many pilot programs suggests it as an instructional variable with real potential for providing more direct kinds of learning experiences, and it should prove to be extremely helpful to those institutions with limited opportunities to place students in more continuous kinds of laboratory situations. Much of the exciting work in this area is relatively unknown to large numbers in teacher education (25; 78), but as McClure (51) points out, many of the same kinds of performance outcomes now assumed to reside as the sole property of field experiences may be attained by using simulation. Certainly the potential for the use of simulation in both admissions and selective retention needs to be explored further. Part of the problem with simulation seems to be the lack of performance criteria with which it can be related.

A supportable viewpoint is that through the selection and organization of instructional techniques students in a program can be provided with a more objective profile of themselves in teaching. It would seem safe to infer that such a profile has the potential to improve the ability of both students and institutions to treat matters of selective retention more completely.

Program influence. As was indicated earlier, Cohen (22), and Ort (60) found positive correlations between descriptions of student teachers, based on the student teaching experience, and future success in teaching. One may infer that a greater emphasis should be placed on the active involvement of students in a laboratory approach. Certainly the conclusions reflected by Bush and Gage (17) and those goals identified by Peck and Brown (61) would lend further support to giving students a more direct contact with teaching and a view of themselves in the teaching act. The popular response to micro-teaching, both as an instructional influence and as the core of a program in teacher education, lends additional support to the potential of making students active partners in their own development (2; 17).

Perhaps of equal significance is the promise of a number of special projects in elementary teacher education. For example, the program described by the University of Georgia (79) provides some 2000 specifications of performance in teaching, and each one, or combinations of specifications, go to make up the basis for judging competence. This same program creates four levels of development through which candidates must pass. This provides four distinct opportunities for the retention or dismissal decisions to be made. Each level is described in terms of performance specifications. The emphasis placed on self-pacing in the Syracuse program (76) and the University of Massachusetts Program (80) provides an additional opportunity for viewing candidate development. Projected far enough, self-pacing could make it possible for a given student teacher to select himself in or out of a program. Other pilot programs such as the one at the University of Pittsburgh (81) make use of a clinical team to work with the individual student to identify strengths and weaknesses as they relate to continuing in the program. Both the Teachers College, Columbia University (77), and the Florida State University (33) programs emphasize the formation and development of extensive data banks for this purpose.

Two important features stand out regarding program influence and selection retention: (a) the primary emphasis is placed on the establishment and use of performance criteria in all programs which it would seem can only strengthen both the objectivity and utility of the program as an influence in selective retention; and (b) the emphasis given to field and laboratory aspects of programs (56) would seem to further enhance the potential for more effective selective retention procedures in programs in teacher education.

Behavior Modification. The notion of behavior modification is raised because of the implications for both admissions and selective retention procedures. Stated simply, if it is true--as studies would suggest (74; 59; 38)--that attitudes toward teaching, interpersonal problems, and the like can be modified, then why become concerned about either admissions or selective retention? Part of the answer resides in decisions made on selection ratios and the development of useful definitions of teacher competence.

Those especially interested in behavior modification may wish to follow the development of the pilot program at Michigan State University (54), where the basic program orientation is built around concepts related to behavior modification. Combining behavior modification programs with group and individual counseling tends to give some hope in our efforts to develop more and better classroom teachers.

Selective retention, then, suffers from many of the same problems encountered in admissions work. There is an absence of criteria and related information from which one may draw conclusions with any degree of confidence. While some students do select themselves out of a program, the reasons vary. There is no reason to assume that such people would not select themselves out of any college or university program. Lurking not far behind all of the more obvious concerns is the basic human concern that these people might have--with more help--remained in the program and might have emerged as highly capable teachers. There are instructional influences--simulation, microteaching, observational learning--that suggest a potential for future development in selective retention, but not enough is known as yet of the potential of such techniques on a broad scale. Program developments also give every indication of providing increased potential to more completely operationalize a process of selective retention. In particular the emphasis placed on the development of performance specifications for learning outcomes suggests a real potential for more objectively treating the process of selective retention. Then, too, instructional techniques and counseling methods hold promise of producing desirable and useful kinds of behavioral modifications. They, in turn, relate directly to selective retention. Currently, however, selective retention processes seem neither to be uniform in development nor extensive in application. In those cases where selective retention processes do exist, the kinds of information used are often incomplete or only distantly related to the effort being made to select and reject candidates already in a program.

Student Involvement

Student involvement has been described as a process of interaction between students and the institution to develop new programs and evaluate existing programs. Implicit is the notion of (a) procedures for involving students, (b) extent and nature of student contributions, and (c) a consideration of the kinds of students involved in program development and evaluation.

Procedures. There is little if any information which suggests that involvement of students achieves any specified goal. At least three different kinds of involvement procedures exist among the institutions: (a) student membership on standing or ad hoc committees as program development and evaluation of existing programs, (b) student membership on student advisory councils, (c) written or oral student reactions to the substance and form of existing programs obtained from students deliberately or on an optional basis by individual faculty. One or more of the above might be defended philosophically on the grounds of student partnership in the on-going enterprise of education. It is simple logic to obtain feedback from those most directly affected by the program. Further, to the extent that one supports a move toward individualization or resolving concerns identified by Fuller (34; 35), then to that extent does student involvement seem defensible. Little if any evidence exists to either confirm or deny the value of making such a provision--or that can help us to select one form of student involvement over others. Put another way, involving students may make sense because it is logically defensible, but in practice there is little evidence to support or reject procedures in terms of differences they produce in the quality of programs developed.

There seems to be more promise in procedures related to experimental programs--career seminars, group counseling, and individual conferences as described in some of the pilot programs (54; 76; 77; 79; 81)--than there is any real evidence already in the literature to support or reject different ways of involving students.

Contributions of students. There is little if any substantial evidence to help in identifying the nature and extent of student contributions to program development and evaluation, or in describing the consequences of such contributions. There is evidence that some of the pilot projects (81) were revised as a result of student response to the original program. There is some defensible basis for inferring that direct student involvement and contributions at Stanford (17) have helped shape the present nature of their program. No direct study can be found in either contributions made or consequences produced from the involvement of students in either program development or evaluation.

Student contributions seem to have been more influential in reshaping and directing the efforts of individual faculty members than in development or evaluation of total programs. While institutions around the country have attempted to support evaluation of instruction through various approaches, thus far the literature does not seem to reflect any effort to systematically study the extent and nature of such contributions or consequences produced.

Students to be involved. At first glance, it would seem logical to assume that the involvement and subsequent contributions of all shapes and sizes of students has much to recommend it. There is some reason to believe, however, that institutions and faculty really should be more discriminating in determining which students are to be involved and the degree of importance to be attached to student contributions.

Carter (21) studied the effect of student characteristics on three different kinds of student evaluations of university instruction. While there were no direct companion studies that either confirm or deny Carter's findings, there is a need for caution on the part of those people seeking to involve students indiscriminately. Carter says that we need to know much more about what kinds of students make what kinds of responses in what kinds of evaluative situations. Only then can we start using student contributions as reflections of thoughtful, sincere, objective, and purposeful responses to either the formation of new programs or the evaluation of existing ones. At this point, it seems that what you find out depends on whom you talk to.

It is one thing to say that students should be involved with existing program evaluators and with those developing new programs. It is a much different thing to suggest the way in which such students involvement is to be structured, what contributions to expect, and which student responses should be given the greatest amount of consideration. Perhaps John Locke gave us the best description for the current level of understanding in matters of student interaction when he refers to knowledge as an "unknown somewhat." What we know about student interaction seems to be an "unknown somewhat."

RECOMMENDATIONS FOR FURTHER STUDY

The recommendations presented now are by-products of the inferential and judgmental activities used earlier. This does not mean that there have not been significant studies with meaning for the standards and their evolution. It does mean, however, that there have not been nearly enough studies done to warrant placing a blind faith in the standards at this time. A real need exists for additional study. The following recommendations, therefore, are made in the belief that further study in these areas can provide a more useful and defensible basis for the determination of standards for admitting, selectively retaining, and actively involving students in program development and evaluation.

Recommendations

1. To accept standards on the basis of some historical notion that opts for their presence because they have always been around is one thing, but to insure the need for them through systematic study is something else. While it may be heretical to suggest it, it seems true that teacher education has not as yet demonstrated directly a need for standards for admissions. A series of longitudinal studies should be initiated by several institutions as a joint enterprise to determine to what extent the presence or absence of such standards makes any real difference in the quality of the candidate produced by a teacher education program.

2. Implicit in the literature related to the standards under consideration here is the belief that until we are able to describe competency or effectiveness in teaching we will never be able to develop useful standards for admissions, selective retention, and programs for student involvement. There is a need, however, for greater emphasis to be given to the matter of defining teacher competence in terms of proximate rather than ultimate outcomes. An increased amount of effort should be put forth to contrast and compare the outcomes produced by preparing teachers using both kinds of outcomes. Perhaps the ultimate decision will require some philosophical deliberations, but the implications for the standards here are quite real. To admit on the basis of potential reflected in terms of one's ability to eventually behave in an ultimate sense may make a considerable difference when compared with potential sought in candidates in terms of proximate outcomes based on a description of learner needs.

3. The matter of priorities and the establishment of priorities in relation to the standards involved here should be given study with attention directed toward the concept of critical functions as they relate to the determination of priorities. It seems evident that priorities now exist and operate in various ways to influence the standards. What is needed is more proof of the consequences produced by establishing different priority rankings in matters related to this section of the standards.

4. The effectiveness of standards that simply prescribe activities in which an institution should engage should be compared with standards that not only prescribe activities but provide known quantities or degrees to be produced by such activities. Such comparisons may be done in terms of the significance of the findings for the attainment of identified program objectives.

5. Attention should be given to the matter of the degree of significance that findings must have before they will be considered as useful for the development of standards. Are we in search of perfect or near perfect correlations or are we willing to gamble that findings of less significance are useful and must be incorporated because the current state of affairs requires action now?

6. Exploratory studies into differences produced by taking individual expectations versus institutional expectations should be initiated. For purposes of the standards reviewed here, making the individual the focus of the preparatory program would seem to imply a much different approach than one built from a primary concern for institutional economy and efficiency.

7. Additional energy should be expended to insure that grades, language competency, emotional stability, and other common criteria have not suffered from a lack of effective means of measuring their relationship to competency.

8. Growing out of recommendation seven, much more emphasis should be placed on building useful evaluative components in the areas of admissions and selective retention. It seems that we have been able

to identify some broad classifications of characteristics and traits with some reasonable potential for predicting future success in teaching. Missing are the necessary techniques to factor such categories and measure the presence or absence of specific behaviors within any degree of validity and reliability. (It is strongly recommended that the several institutions give institutional recognition at the doctoral level to the development of hardware as a legitimate form of doctoral study.)

9. Longitudinal studies should be initiated involving variables in admission that are quite removed from the conventional criteria of grade point average, grades in selected courses, and language competency. It is recommended for example that Cook's (23) work on personal data be pursued further, and that the work initiated by Lucio and others (46) in matters of psychophysiological relationships to potential effectiveness be projected.

10. The profession and respective institutions should encourage longitudinal studies reflecting testing of hypotheses both in laboratory and in field settings. The implications of such an emphasis are clearly understood in terms of institutional personnel policies involving "publish or perish," but as several writers have pointed out, part of our problem stems from a lack of extensive study over a long period of time to permit us to say with any degree of certainty what the implications of the findings produced might be. One cannot be preoccupied with gaining journalistic visibility and still focus primary attention on long-range investigations.

11. The institutions subscribing to the AACTE should be willing to invest some institutional funds into the investigation and subsequent production of findings in this area of the standards. This implies money; and it suggests institutional endorsement and support for variations on programs, program components, and staff utilization.

12. The whole area of student concerns should be investigated as it relates to admissions, selective retention, and student involvement. Fuller's work (34; 35) and the separate work of Muro (55) point to a real need to give serious consideration to the implications of student concerns about becoming a teacher. In connection with the efforts to establish concerns within the area of teacher education, Fuller (35) identifies several different questions in need of study including,

    (a) Are concerns really related to teacher behavior?
    (b) Are concerns manipulable?
    (c) What tasks and competencies are involved in different levels (lower-higher) stages of concerns?
    (d) Is a concern a function of person, situation, or both (35)?

13. Explorations into instructional and counseling techniques in behavior modification should be expanded and intensified as useful adjuncts to admissions, selective retention, and student involvement. Preliminary efforts seem to support and justify a major effort in this direction. Positive findings could have great implications for the

entire matter of admissions, individualization of programs and selective retention. Further, positive findings might go a long way toward helping provide a more definitive base for one or more philosophical differences perplexing proponents and opponents of such matters as self-pacing, instructional modules, model imitation, career counseling and so forth.

14. A major effort should be undertaken to explore the feasibility and relevance of performance screening as a major tool in admissions and selective retention in teacher education. A lack of commonly agreed to performance specifications may hamper efforts in this matter, but the same lack of agreement on performance specifications, has not allowed to deter a considerable amount of study in other areas such as program development. Performance screening deserves more serious treatment as a viable alternative to the use of conventional criteria than it would seem it has thus far been accorded.

15. Investigations into various kinds of data collection procedures should be initiated to determine (a) which procedures produce the most useful kinds of data in terms of admissions and selective retention, and (b) which procedures can realistically be employed within the several institutions in terms of institutional capacity to incorporate such procedures. It is also recommended that this be a longitudinal task.

16. The decision-making process related to admissions and selective retention should be studied to determine (a) whether or not decision-making is automatic once criteria are established; (b) whether the use of a review committee is more effective than a single decision maker such as a director in terms of time, discrimination, consistency and certainty; (c) whether there are other decision-making schemes that can be used, i.e., counseling or conference decisions, and so forth, that are more effective than others now in use. As clarity develops in the area of competency or effectiveness, the scope of criteria for admissions will probably expand and both continuous and dichotomous data will be used. As the criteria expand and become more complex, the need to more fully understand and implement an effective decision making process will also become apparent. Work on this ought to begin immediately.

17. Recruitment should be investigated in terms of its feasibility and relatedness to admissions and selective retention. The literature seems to suggest that recruitment never received sufficient study to know whether or not it really had any potential as a part of the admissions and selective retention operations.

18. The nature and extent of the operation of selection ratios in teacher education should be investigated. In an ideal sense, if a recruitment program was effective, the selection ratio could always remain low and the potential competitiveness might foster a more objective system of admissions and retention. At the same time, a selection ratio also would require greater attention to time, discrimination, consistency, and certainty in admissions and selective retention.

19. Additional investigations into the area of withdrawals from programs should be projected. Such information has relevance for the evaluation of existing admissions criteria and for the development of future criteria. It is further recommended that such investigations be maintained on a continuous basis and some yearly accounting for all institutions be initiated. A general lack of data banks would hamper the ability of the institutions, standards or not, to really be able to systematically evaluate their own efforts. The investigation of the nature and extent of withdrawals and continuations can provide useful data for program decisions.

20. The concept of self-selection in terms of admissions and selective retention in programs should be studied. It is argued by some that students do select themselves in and out of the program, but neither the nature nor extent of such activity is understood or adequately described. Here again, such information does not either support or refute any standard.

21. An additional study should be promoted in the area of effects produced by variations in instructional techniques in education programs. In particular, such studies ought to be concerned with differences produced in the kinds of students staying in or dropping out of programs. These same studies could also determine whether or not different and more complete kinds of information related to selective retention might not be produced. Perhaps the time has come for education to study the impact and significance of variations on its own instructional theme after studying so intensively the relationship between various instructional techniques and outcomes at other educational levels.

22. The use of instructional modules as organizing centers for programs should be studied in connection with admissions, selective retention, and student involvement. What differences do the use of instructional modules produce when compared with the use of conventional approaches? To what extent are similar outcomes achieved by two or more differing programs? Is self pacing or individualized instruction as a program approach a viable alternative to conventional patterns in terms of more effectively achieving designated outcomes? How does the concept of differentiated staffing fit in as a model for program development? How would a program organized around the concept of differentiated staffing be similar to or different from conventional programs, and how might the two compare in their ability to achieve designated learning outcomes in a program of teacher education? Is career counseling a legitimate component of program development and what differences are produced as a result of its presence or absence in programs of teacher education? Does its presence or absence affect the kinds of pupils retained in or dropped from the program? What differences are produced when programs are solely field oriented as opposed to programs partially field and partially classroom oriented as opposed to programs solely classroom oriented? No doubt the information produced by the evaluation studies in conjunction with the pilot programs in Syracuse University (76), Florida State University (33), University of Toledo (82) and others will produce valuable information related to all aspects of the standards under consideration.

23. Simulation should be studied more extensively to determine its implications for use in admissions and selective retention.

24. Additional attention should be given to role expectations of students seeking admissions, and that the relative influences of both student and staff expectations be investigated in terms of implications for all aspects of the standards under consideration.

25. The role of counseling should be studied in terms of its contribution both to admissions and selective retention. It is also recommended that it be studied in terms of its involvement in the formation of a basic approach to teacher education.

26. The nature and extent of student contributions to program development and evaluation should be studied. It might be helpful to know what the differences would be between programs developed conventionally and a teacher education program planned cooperatively by students and faculty. The similarities and differences produced might have real implications for program development and evaluation. Even a status study into the extent and nature of student contributions to program development and evaluation would be helpful in shaping additional investigations.

27. A study should be initiated involving the procedures available for use in involving students in program development and evaluation. Procedures and contributions are probably related, but there is a need to know what differences result from the use of varying procedures for involving students.

28. Efforts should be made to determine what differences, if any, are produced by systematic and deliberate involvements of students in program development and evaluation. This is fundamental, of course, to the whole question of whether or not student involvement in program development and evaluation really makes any difference at all.

29. A study should be made of the nature of students who are or who could be involved in a system of program development and evaluation. Who are the students whose involvement produces the greatest contributions? How do students differ in terms of responses to different procedures or different kinds of contributions made? Is student involvement a matter of personal responsibility or should institutions actively seek student members for participation in such tasks? On what grounds could distinctions be made between student members if such distinctions proved to be desirable?

A BEGINNING

It seems appropriate to think in terms of a beginning from this point. The standards, it has been concluded, exist more at the pleasure of conventional wisdom than any real empirical base. The state of the research base is such that much of it is not directly related to the standards necessitating inferential and judgmental processes as a means of demonstrating support. The need is to begin to study the standards

directly, to establish dependent and independent variables suggested by the standards and to conduct related research suggested by such variables, and to develop standards that are both defensible in terms of contributions they make to improving or insuring a minimally desirable level of accomplishment by the institutions and that make comparisons a reality. There is a need for member institutions to begin to work cooperatively, not competitively, for the betterment of programs. The time has come for institutions to come of age and put common good above institutional prestige. As regards the standards and their research base at this point in time, then, where else is there to begin except at the beginning?

Bibliography

1. Aiken, L. R. "Three Alignment Charts for Use in Selective Admission," *Journal of Educational Research*, 62:14-17; September 1968.

2. Allen, Dwight W. "Micro-Teaching: A New Framework for In-Service Education," *High School Journal*, 69:355-62; May 1966.

3. Altenheim, M. R. "Guidance and Selection in a Teacher Education Program: Hunter College of the City University of New York," *The Journal of Educational Sociology*, 36:70-78; October 1962.

4. American Association of Colleges for Teacher Education. "Evaluation and Selection Instruments in Teacher Education," *Action for Improvement of Teacher Education*. Eighteenth Yearbook. Washington, D.C.: the Association, 1965.

5. Barr, Arvil S. "The Nature of the Problem," *Journal of Experimental Education*, 30:5-9; September 1961.

6. ---. "Teaching Effectiveness and Its Correlates," *Journal of Experimental Education*, 30:134-56; September 1961.

7. ---. "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations," *Journal of Experimental Education*, 30:1-156; September 1961.

8. Beeley, A. J. "Clinical Techniques for the Selection of Prospective Teachers," *School and Society*, 50:183-85; August 1939.

9. Biber, Barbara, and Charlotte B. Winsor. "An Analysis of the Guidance Function in a Graduate Teacher Education Program," *Mental Health and Teacher Education*. Forty-Sixth Yearbook. Washington, D.C.: Association for Student Teaching, 1967.

10. Bledsoe, Joseph C. *Comparison Between Selected Characteristics and Performances of Provisional and Professionally Certified Beginning Teachers in Georgia*. Final Report. Athens: University of Georgia, 1967.

11. Bolton, Dale L. *Variables Affecting Decision-Making in the Selection of Teachers*. Seattle: University of Washington, 1968.

12. Bondi, Joseph C. "Feedback in the Form of Printed Interaction Analysis Matrices As a Technique for Training Student Teachers." Paper presented at American Educational Research Association annual meeting, 1969, Los Angeles.

13. Bowman, David L. *Quantitative and Qualitative Effects of Revised Selection and Training Procedures in the Education of Teachers of the Culturally Disadvantaged*. Final Report. Oshkosh: Wisconsin State University, 1968.

14. Boxhill, O. J. "Scale to Aid in the Retention and Dismissal Decision," *Personnel and Guidance Journal*, 45:53-55; 1966.

15. Bradley, R. C. "Clarifying the Supervising Teachers Role," *The Teachers College Journal*, 38:92-94; 1966.

16. Bunyan, Lorne W. "A Comparative Investigation of the Effects of Three Different Types of Supervisory Evaluations on Certain Latent Functions of Student Teaching." Unpublished Doctor's dissertation, University of Oregon, 1969.

17. Bush, Robert N., and N. E. Gage. "Center for Research and Development in Teaching," *Journal of Research and Development in Education*, 6:85-105; Summer 1968.

18. Caldwell, M. S. "Student Selection, Risk, and the Error Continuum," *Peabody Journal of Education*, 44:230-35; January 1967.

19. Canadian Teachers' Federation. *Bibliographies in Education, Number One: Pre-Service Teacher Education in Canada*. Ottawa, Ontario: the Federation, 1969.

20. Cardozier, V. R. *Undergraduate Academic Achievement and Teaching Performance*. College Park: University of Maryland, Maryland Agriculture Experimental Station, 1965.

21. Carter, Robert E. "The Effect of Student Characteristics on Three Student Evaluations of University Instruction." Unpublished Doctor's dissertation, Indiana University, 1968.

22. Cohen, Arthur M., and Florence B. Brawer. "Adaptive Potential and First Year Teaching Success," *Journal of Teacher Education*, 18:179-85; 1967.

23. Cook, Desmond L. "The Personal Data Form As a Predictor of Success in Teacher Education Programs and Entry into Teaching," *Journal of Teacher Education*, 15:61-66; March 1964.

24. Cornett, J. D. "Effectiveness of Three Selective Admissions Criteria in Predicting Performance of First Year Teachers," *Journal of Educational Research*, 62:247-50; 1969.

25. Cruickshank, Donald R., and Frank W. Broadbent. *Simulation in Preparing School Personnel*. Washington, D.C. ERIC Clearinghouse on Teacher Education, November 1969.

26. DiTosto, Evelyn. "The Factors of Interpersonal Relations in Compatibility to Productivity of Student Teacher-Supervisor Dyads." Unpublished Doctor's dissertation, West Virginia University, 1968.

27. Domas, S. J., and D. Tiedeman. "Teacher Competence: An Annotated Bibliography," *Journal of Experimental Education*, 19:101-218 December 1950.

28. Durflinger, Glenn W. "Recruitment and Selection of Prospective Elementary and Secondary School Teachers," *Review of Educational Research*, 33:355-68; October 1963.

29. Edson, William H., and C. J. Glatzbach. "Recruitment and Selective Admissions of the Prospective Teacher," *Guidance in Teacher Education*. Thirty-Sixth Yearbook. Washington, D.C.: Association for Student Teaching, 1957.

30. ---, and Don Davies. "Selectivity in Teacher Education," *Journal of Teacher Education*, 11:327-34; September 1960.

31. Fisher, Rheua Dale S. "A Study of Grades and Test Scores in a Selective Admissions Program for Teacher Education." Unpublished Doctor's dissertation, Oklahoma State University, 1968.

32. Fliger, R. J., and Karl Dowing. "Grades and Teacher Education," *Illinois Education*, 56:305-07; March 1968.

33. Florida State University. *Summary of a Model for the Preparation of Elementary School Teachers*. Tallahassee: the University, 1968.

34. Fuller, Frances F. "Concerns of Teachers: A Developmental Conceptualization," *American Educat'onal Research Journal*, 6: 207-26; March 1969.

35. ---, Geneva H. Pilgrim, and Alma M. Freeland. "Intensive Individualization of Teacher Preparation," *Mental Health and Teacher Education*. Forty-Sixth Yearbook. Washington, D.C.: Association for Student Teaching, 1967.

36. Hart, James E. "Attitudinal Change in Pre-Service Teacher Education." Unpublished Doctor's dissertation, Ball State University, 1968.

37. Hayes, Allen P. "Effects of Cooperating Teachers and College Supervisors on Student Teachers' Beliefs." Unpublished Doctor's dissertation, University of Florida, 1968.

38. Howard, Rosalie D. "The Effects of Structured Laboratory Learning on the Interpersonal Behavior of College Students." Unpublished Doctor's dissertation, University of Oregon, 1969.

39. Hudgins, Bryce B., and William P. Ahlbrand, Jr. *A Study of Classroom Interaction and Thinking*. Interim Report. St. Ann, Mo.: Central Midwestern Regional Educational Laboratory [1967].

40. Husemoller, Kenneth E. "The Prediction of Freshman Academic Success at Eastern New Mexico University, Roswell, By Means of Selected Demographic and Standardized Test Data." Unpublished Doctor's Dissertation, Colorado State College, 1969.

41. Ivanoff, J. M., and others. "Use of Intellective and Non-Intellective Factors in Classifying Female Elementary and Secondary Teacher Trainees," *Journal of Experimental Education*, 34:55-61; 1966.

42. Joyce, Bruce R., and Richard E. Hodges. "The Use of Developmental Studies of Teaching Styles for Research on Teacher Education." Paper presented at American Educational Research Association annual meeting, 1966, Chicago. (Mimeographed)

43. Koran, Mary Lou. "The Effects of Individual Differences on Observational Learning in the Acquisition of a Teaching Skill." Unpublished Doctor's dissertation, Stanford University, 1969.

44. Lindsey, Margaret, and others. *Annotated Bibliography of the Professional Education of Teachers*. Washington, D.C.: Association for Student Teaching, 1969.

45. Loftis, Helen A. "The Study of Commitment to Teaching," *Journal of Home Economics*, 56:157-63; March 1964.

46. Lucio, William H., Marion A. Wenger, and Thomas D. Cullen. *Psychophysiological Correlates of Female Teacher Behavior and Emotional Stability: A Seven Year Longitudinal Investigation*. CSEIP Technical Report No. 3. Los Angeles: University of California, November 1967.

47. McGee, Robert M. "Admission-Retention in Teacher Education," *Journal of Teacher Education*, 12:81-85; March 1961.

48. Mascho, Beth, and others. *The Elementary Education Selection Research Project*. Muncie, Ind.: Ball State Univeristy, 1966.

49. Mason, Ward S., Robert J. Dressel, and Robert K. Bain. "Sex Role and the Career Orientations of Beginning Teachers," *Harvard Educational Review*, 29:370-83; Fall 1959.

50. McCall, Kendrick M. "The Modification of Teacher Behavior and Self-Perceptions: Effects of Dissonance and Coded Feedback." Unpublished Doctor's dissertation, Rugters University, 1969.

51. McClure, L. Morris. "Screening For Teacher Education," *Educational Leadership*, 25:44-48; November 1967.

52. Medley, Donald M. Some Notes on the Research Base for the *"Standards for the Accreditation of Teacher Education Basic Programs: Standards of Resources."* Princeton, N.J.: Educational Testing Service, 1967.

53. Metzner, Seymour. *An Empirical Criterion Validation Study of Some Psychological Findings Related to Elementary School Teachers*. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education, 1967.

54. Michigan State University. *Summary of Behavioral Science Elementary Teacher Education Programs*. East Lansing: the University, 1968.

55. Muro, J. J., and G. M. Denton. "Expressed Concerns of Teacher Education Students in Counseling Groups," *Journal of Teacher Education*, 19: 465-70; Winter 1968.

56. Northwest Regional Educational Laboratory. *Summary of a Competency Based, Field Centered, Systems Approach to Elementary Teacher Education*. Eugene, Oregon: the Laboratory, 1968.

57. Notestine, Earl B. "A Comparative Study of Student Retention and Withdrawal in the School of Humanities, Social Science, and Education at Purdue University." Unpublished Doctor's dissertation, Purdue University, 1969.

58. O'Donnell, Richard W. "Proposed Revision of the Admission to Teacher Education Policy of the University of Maryland." Unpublished Doctor's dissertation, University of Maryland, 1968.

59. O'Reilly, Ruth H. "Use of Differential Counseling Methods in Helping Student Teachers Cope with Interpersonal Problems." Unpublished Doctor's dissertation, University of California at Los Angeles, 1968.

60. Ort, Vergil K. "A Study of Some Techniques Used for Predicting the Success of Teachers," *Journal of Teacher Education*, 15:67-71; March 1964.

61. Peck, Robert F., and Oliver H. Brown. "Research and Development Center for Teacher Education," *Journal of Research and Development in Education*, 1:106-24; Summer 1968.

62. Pierson, Robert A. *The G. P. A. Criterion and Selective Retention in Teacher Education* [Macomb: Western Illinois University], 1966.

63. Rabinovitz, W., and H. E. Mitzel. "Some Observations on the Selection of Student for Teacher Education Programs," *Journal of Teacher Education*, 12:157-64; June 1961.

64. Ray, Elizabeth M. *An Experimental Attempt to Maximize the Professional Potential of Home Economics Teachers Through a Program of Group Counseling in College*. University Park: Pennsylvania State University, 1968.

65. Rossman, J. E. "Experimental Study of Faculty Advising," *Personnel and Guidance Journal*, 46:160-64; October 1967.

66. Rousseau, Leon, Herbert Hite, and William Drummond. "Appendix J. An Experimental Model to Enable Instructional Managers To Demonstrate Interaction Competency," *A Competency Based, Field Centered, Systems Approach to Elementary Education*. Final Report. Vol II. Portland, Ore.: Northwest Regional Educational Laboratory, 1968.

67. Ryans, David G. *Characteristics of Teachers*. Washington, D.C.: American Council on Education, 1960.

68. Sharie, Abdullah N. "A Comparative Study of Belief Systems of College Students Majoring in Different Fields." Unpublished Doctor's dissertation, The University of Oklahoma, 1969.

69. Sharon, Amiel Thomas. "The Effect of Instructional Conditions in Producing Leniency." Unpublished Doctor's dissertation, University of Maryland, 1968.

70. Shaver, James P., and Hyrum E. Richards. *Open-Closed Mindedness of College Students in Teacher Education*. Logan: Utah State University, Bureau of Educational Research, 1968.

71. Slaughter, C. H. "Proposed Screening Program for Elementary Teacher Candidates," *Journal of Teacher Education*, 20:343-46; Fall 1969.

72. Smith, G. "Admissions Problems at a University School of Education," *Journal of Educational Sociology*, 25:325-28; March 1962.

73. Sowards, G. W. *A Model for the Preparation of Elementary School Teachers*. Tallahassee: Florida State University, 1968.

74. Spaights, Earnest, and L. P. Blum. "Social Reinforcement Counseling as an Adjunct to a Course in Education," *Journal of Educational Research*, 60:256-59; February 1967.

75. Stout, Ruth A. "Selective Admissions and Retention Practices in Teacher Education," Part III, *Journal of Teacher Education*, 7:422-32; December 1957.

76. Syracuse University. *Summary of the Specifications for a Comprehensive Undergraduate and Inservice Teacher Education Program for Elementary Teachers*. Syracuse: the University, 1968.

77. Teachers College, Columbia University. *Summary of the Teacher-Innovator: A Program to Prepare Teachers*. New York: the College, 1968.

78. Twelker, Paul A. *Simulation: An Overview*. Monmouth, Ore.: Teaching Research Division, Oregon System for Higher Education, 1968. (Mimeographed).

79. University of Georgia. *Summary of the Georgia Educational Model Specifications for the Preparation of Elementary Teachers*. Athens: the University, 1968.

80. University of Massachusetts. *Summary of a Proposed New Program for Elementary Teacher Education*. Amherst: the University, 1968.

81. University of Pittsburgh. *Summary of a Model of Teacher Training for the Individualization of Instruction*. Pittsburgh: the University, 1968.

82. University of Toledo. *Summary of the Educational Specifications for a Comprehensive Elementary Teacher Education Program*. Toledo, Ohio: the University, 1968.

83. White, Kinnard. "Social Background Variables Related to Career Commitment of Women Teachers," *Personnel and Guidance Journal*, 45:648-52;: March 1967.

84. Yee, Albert H. *The Student-Teaching Triad--The Relationship of Attitudes Among Student Teachers, College Supervisors, and Cooperating Teachers*. Austin: University of Texas, College of Education, 1967.

RESEARCH RELATED TO THE EVALUATION OF GRADUATES OF TEACHER PREPARATION PROGRAMS

R. L. R. Overing
University of British Columbia

The fifth of categories in the *Recommended Standards for Teacher Education* (1969) has to do with standards for evaluation of graduates, program review, and long-range planning. The review which follows deals only with the topic "evaluation of graduates" and does not touch upon either the topics of program review or long-range planning.

This review deals with studies of the following sorts:

1. Studies evaluating graduates of teacher preparation programs against internal criteria (i.e., using as criteria outcomes of the program specified in terms of teacher behaviors and characteristics).
2. Studies evaluating graduates of programs against external criteria (i.e., using pupil change as a criterion for the evaluation of the graduates' effectiveness).

Studies which attempt to determine experimentally which characteristics of a teacher or a teaching situation interacted with particular learner characteristics to facilitate or inhibit learning. Nor have I included studies which concern themselves with the development and validation of instruments designed to record teacher behavior or characteristics. Finally, I have not included studies that are concerned solely with developing predictors of success in the training program or in subsequent teaching.

I have not excluded studies of these sorts because they are of no concern or of secondary importance. On the contrary, I consider such studies to be of fundamental importance in that, ultimately, they must provide both the empirical basis on which we build programs of teacher education, and the instrumentation for selecting and for evaluating our graduates. Such studies, however, must be omitted from this review for it is not my charge at this time to consider directly the vast area of research on teacher effectiveness.

My concern, as I have attempted to make clear above, is to search for studies where (a) either specific objectives were formulated and a serious attempt was made to evaluate the program and/or the graduates, using the objectives as criteria, or (b) the graduates of a program were evaluated using the achievement of their pupils as a criterion. Using criteria such as these it is obvious that I will also omit from this review any mention of evaluation studies which employ as their principal source of data the opinions of the graduates of the program. Studies of this sort seem frequently to be conducted by institutions engaged in training teachers. The ones that I have encountered strike me as being of little use as a main source of data for future decision making.

I have restricted the range of studies reviewed to the two classes noted above for two principal reasons. First, I believe that evaluation studies of these sorts, rigorously pursued, are the ones most likely to advance both our understanding of the nature of an effective training program and of our knowledge of the technology necessary to design, to describe and to evaluate improved training programs. That such studies are likely to be uncommon is suggested by Stiles and Parker who state, "Evaluation of entire teacher education programs, or even of segments of programs, is spotty and inadequate" (23: 1418). Second, the *Recommended Standards* (3) themselves identify evaluation studies of these sorts as the ones which they recommend and hope to promote. The authors state, "The ultimate criterion for judging a teacher education program is whether it produces competent graduates who enter the profession and perform effectively" (3: 12). And a few lines further, they state, "Any effort to assess the quality of the graduates requires that evaluations be made in relation to the objectives sought. Therefore, institutions use the stated objectives of their teacher education programs as a basis for evaluating the teachers they prepare" (3: 12).

These two statements makes it apparent that two quite different criteria are being advocated, and we know that it is quite possible that these two criteria may be independent of one another. That is to say, the stated objectives of the teacher education program may bear no relationship to effective teaching. Hopefully, this is not so. Nevertheless, we must always ask of any program that specifies its objectives, "What are the grounds for these objectives? Which objectives have a hypothetical basis, which have an analytic basis, and which have an empirical basis?" Questions of this sort relate to problems of criteria. The fact that I have restricted my remarks to two kinds of criteria, namely (a) specified teacher behaviors and characteristics and (b) pupil change, does not mean that there are not, or cannot be, other criteria. This whole problem of criteria is obviously of fundamental concern in any attempt to evaluate graduates of teacher training programs. Nobody should embark on such a venture without being thoroughly familiar with at least, the reports of AREA (1; 2), Rabinowitz and Travers (17), Morsh and Wilder (14), Mitzel (13), Ryans (18; 19), Barr (4)--all of which attempt to cast some light on this criteria problem.

In passing it may be of interest to note that my search of the literature was made principally within the ERIC indexes, 1965-68, and within Psychological Abstracts, 1960-68. The rubrics ("descriptors") used within ERIC were Evaluation, Evaluation Criteria, Evaluation Methods, Evaluation Needs, Evaluation Techniques, Program Evaluation, Teacher Evaluation, Teacher Proficiency, Teacher Rating, Effective Teaching, Teacher Education Curriculum, Teacher Behavior, Teacher Certification, Task-Performance, Observation, Behavior Change, Professional Education, Professional Training, Objectives, Measurement Techniques, and Preservice Education.

The rubrics used within Psychological Abstracts were Job Evaluation, Evaluation, Teacher Training, and Training.

In all I followed up some 200 references, which from their titles seemed appropriate. That the elephant labored and gave forth a mouse will quickly become apparent as I read on.

STUDIES EVALUATING GRADUATES OF TEACHER PREPARATION PROGRAMS AGAINST INTERNAL CRITERIA

Evaluation studies of graduates of teacher preparation programs which use specified objectives of the program as criteria require two general components, viz., first, a set of specified objectives describing the abilities, the characteristics and dispositions which graduates of the program are expected to exhibit; second, a set of instruments and techniques for measuring the extent to which graduates of the program exhibit these abilities, characteristics and dispositions. To the extent that we may also wish to say that the abilities, characteristics and dispositions. To the extent that we may also wish to say that the abilities, characteristics and dispositions exhibited by the graduates are due to the effects of the program we will also have a set of instruments and techniques to obtain pre-measures of these same graduates when they entered the program. But that is a slightly different question which need not concern us directly here. However, we should keep in mind the evaluations of program effectiveness as contrasted with evaluation of graduates of this program may have to use this pre-test, programming, post-test model.

Large-scale studies which actually have attempted to determine the extent to which graduates of a teacher preparation program have acquired the behaviors and characteristics described in the program objectives are rare, and in any pure form, seem to be non-existent. Their frequency may increase, however, for the recent USOE-funded Comprehensive Elementary Teacher Education Models have all been formulated around the central idea of specified teacher competencies (8). For example, Dickson and others (7) have listed 818 program objectives, formulated in terms of specific teacher behaviors and, in what is frequently a very general manner, have also described how participants in the program will be evaluated to determine if they have met the criteria. The description of the evaluation techniques is general in the sense that frequently there is no mention of the specific instruments and techniques by which the evaluation will be carried out. As the design and validation of such instruments is normally a demanding, lengthy, and expensive task we should recognize the significance of this lack of specificity. Nevertheless, the availability of teacher preparation programs built around specified objectives presumably means that the attempt will now be made to evaluate the extent to which these objectives have been attained.

One study, though by no means a model, may suggest something of the state of the art and of the problems still to be solved. The Final Progress Report of a Ford Foundation-sponsored teacher education project carried out at the University of Missouri at Kansas City (published in 1967) is devoted largely to evaluation. While evaluation of several sorts was attempted, only those parts of the evaluation study which concerns themselves with certain pre-specified verbal behaviors of the graduates approximate the type of evaluation study here under review.

Graduates of the program were evaluated during their first year of teaching to see if their teaching behaviors reflected the specific objectives of the part of their program which had dealt with the teaching of cognitive behaviors. This program component had attempted to train them to teach so as to give particular emphasis to higher level behaviors. Specifically, audio-tapes were made of two lessons for each of a group of experimental teachers and each of a group of controls (total N = 40). These tapes were than analyzed to determine (a) the percent of teacher verbal behavior which fell into each of Bloom's categories for the cognitive domain, and (b) the number of pupil responses induced by teacher questions. No significant differences were found. However, when the experimental group was divided in two, to form a group with high academic achievement and a group with low academic achievement, significant differences between certain of the sub-groups of these high and low groups emerged, favoring the high group. With the exception then of these small sub-groups, there was no evidence that graduates of the program were teaching in a manner to reflect the objective criteria of the program. Whether the n.s.d. results are due to lack of treatment difference or to reliability and sampling problems is not apparent.

While I was able to locate no other large scale studies which attempted to evaluate their graduates against internal criteria, there are two studies which I would like to mention in this section. In both cases the behavior of the graduates of the program was measured, but in neither case were there explicit pre-specified program objectives against which the behavior measured could be evaluated. Sandefur and others (20) devised an experimental program which attempted (a) to identify and to organize knowledge related to teaching and learning; (b) to design and to implement a series of laboratory experiences; and (c) to evaluate the extent to which teacher behavior was affected. Essentially, they attempted to coordinate laboratory experiences allowing observation and participation with appropriate readings, and to conduct the whole program in a relatively informal, non-threatening seminar context. Sixty-two members of this experimental program were then compared with fifty-two members of a conventional program within the same institution. Data on classroom behavior were collected during student sessions using Ryans' classroom observation record (19: 83-92) and Hough's modification of Flanders' system of interaction analysis. Additional data were collected using student-teaching grades and the National Teachers Examination. Hypotheses looked for differences in teacher behavior, teaching patterns, pupil behavior, student-teaching grades and professional knowledge. In all categories except professional knowledge, as measured by the National Teachers Examinations, student teachers from the experimental group and the pupils under their direction showed significant differences in the direction of behaviors generally held to be desirable. For example, experimental teachers showed significantly more use of behavior which could be categorized as praise, acceptance and use of pupil ideas, student talk, demonstration, and so forth. Their pupils were judged more alert, responsible, initiating, fair, democratic, and so forth. Thus, while no program objectives had been pre-specified, the program designers were prepared to say that the classroom behavior of participants was of the sort which they wished to produce by their program.

In a sense, the "desirable" and the "undesirable" behaviors which the instruments were designed to record provided an implicit set of behaviors to serve as objectives of the program. Obviously, it would be a relatively simple matter to make these objectives explicit. While there may be limitations to this approach, it seems not a bad idea for program designers to concern themselves with behaviors for which there already exist measuring instruments of some demonstrated reliability and validity. Approximately 80 such direct observation instruments and techniques are summarized in the Simon and Boyer anthologies (21; 22).

A second and somewhat similar case is provided by Corle (6) who compared 16 intermediate mathematics teachers who received inservice training via a 15 week ETV program and 16 who did not view the program. Ss were visited seven times before the inservice training began and 23 times during the program. Behavior was recorded on a modification of Medley and Mitzel's OSCAR, designed for elementary mathematics classrooms. Only one behavior category of the six recorded showed a significant difference in favor of the experimental group. Lack of feedback, lack of shaping and short duration of the training program are given as possible reasons for the lack of behavioral change evident. However, while the author had no pre-specified objectives for his program, he was prepared in his discussion section to judge certain of the behavior categories of the OSCAR (EM) as more or less desirable and to imply that his course was successful to the extent that it moved teachers towards these desirable categories. Thus, he, like Sanderur, was using the behavior categories of his instruments as the implicit objectives of his program.

STUDIES EVALUATING GRADUATES OF TEACHER PREPARATION PROGRAMS AGAINST EXTERNAL CRITERIA

I was unable to locate any studies whatsoever which evaluated graduates of a teacher preparation program against the criterion of pupil achievement. Studies attempting to use this "ultimate" criterion of pupil achievement are still small scale and concerned with developing criterion instruments or concerned with mapping teacher behavior in order to identify significant teacher variables. The study which came closest was that reported by Popham and Baker (15; 16). This study attempted to determine if teachers who differed greatly in terms of experiences and training would be differentially effective in promoting pupil change. The underlying purpose was to validate a test of teacher effectiveness, using pupil achievement as a criterion. The study directors, building on the observation of Turner and Fattu (27)--that the relative effectiveness of teachers could be judged only when they were attempting to teach to the same objectives--provided teachers with (a) a set of instructional objectives (à la Mager), (b) suggested a variety of means to teach these objectives, (c) spelled out the subject matter content, and (d) provided a pre- and post-test which the participating teachers neither saw nor administered. In the hope of obtaining differences between teachers, two apparently very different groups were formed, one consisting of trained teachers who (a) had received A in a curriculum and instruction course emphasizing the construction and use of behavioral objectives, (b) had social studies majors and, (c) had been judged superior by their supervisors. The other group was made up of housewives who (a) had had no formal

teaching experience or teacher training, (b) had at least two years of college, and (c) had been enrolled as social studies majors. There were no significant differences whatsoever between the achievement scores of the pupils whether taught by the experienced teachers or the inexperienced teachers. Nor were there any differences in attitudes expressed by the pupils, nor did the teachers themselves differ in their reactions to the materials, the objectives etc. which were provided for them.

Popham suggests that the principal reason explaining why there were no differences in pupil achievement may be that "experienced" teachers are no more experienced than intelligent lay people in bringing about change in pupils. This is not to say that the trained teachers do not possess certain specialized skills and knowledge; rather that this skill and knowledge does not seem to be particularly related to pupil change.

I have dwelt at some length with this study, even though it does not specifically set out to evaluate graduates of a program, for two reasons. First, I have been able to locate so little else to report, and, second, I have wished to emphasize for you the complexity of the problem of evaluation which we are considering. Popham is an extremely imaginative, intelligent researcher who spent a lot of time, and devoted a lot of resources to design a test which would discriminate between teachers. To increase the likelihood of his obtaining differences he took two apparently very different groups of teachers. Despite these efforts he was able to detect no differences. If nothing else this suggests there are no simple-minded easy solutions to the problem of evaluating graduates of programs using pupil achievement as the criterion.

CONCLUSIONS

I am afraid that this paper advances our understanding of the nature and problems of evaluating graduates of teacher preparation programs very little. Perhaps it will be of some use if it brings to our attention the fact that while many writers have advocated the approach to evaluation now suggested in the *Recommended Standards*, almost no one has attempted it. Some writers (for example, Woodruff, 29) believe we are right on the edge of being able to evaluate our products satisfactorily. Woodruff writes, "It is doubtful that we could have taken this direction (i.e., the evaluation of program products) earlier with any realistic chance for success, but I am convinced we can do so now, and indeed that we must for the sake of professional responsibilities" (29: 245). Fattu (8), however, raises the question of whether all components necessary for an invention (in our case the means and technology of product evaluation) are available to the people trying to do the inventing. For example, do we have any reasonably satisfactory set of criterion behaviors around which to design our programs and against which to evaluate our graduates? Dickson and others (7) state: "What a teacher does as he performs his tasks must be determined before the knowledge and experience needed in developing these teaching skills can be ascertained" (7: 90). We need to ask ourselves to what extent the significance of the various teacher behaviors which are offered as program objectives has been empirically determined and to what extent their significance is merely conjectured.

The *Recommended Standards* state that the means now available for making such evaluation (i.e., the evaluation of program products) are not fully adequate. This may turn out to be the understatement of the year. There is no doubt that much rigorous and imaginative basic research is being done in the area of program product evaluation. For example, McGuire (12) writes, in the context of medical education, that

> . . . products of medical education are being studied by systematic evaluation procedures which include: empirical determination of essential components of professional competence, employment of simulation techniques to supplement more conventional methods of assessment, application of pre-established standards, and utilization of numerous feedback mechanisms to assure fuller exploitation of evaluation data. Such evaluation studies are being employed not only to assess individual achievement of critical performance requirements, but also to identify differential rates and patterns of progress toward these goals, to determine the relation between these patterns and important independent variables in the learning situation, to guide curricular development, and to provide evidence of value in redefining the goals themselves (12: 51).

Some of these same kinds of studies, only focusing on teacher education, are undoubtedly being attempted right now. All of them are being advocated. A balanced set of the kinds of studies listed by McGuire, above, actually would contain all the sufficient and necessary components for the evaluation of program graduates. But the very fact that research and developmental-type studies are being undertaken which focus on individual components of the evaluation process, serves to raise the question, "Have we as yet the means and techniques to conduct evaluation of teacher preparation programs of the sort advocated in the *Recommended Standards*?" My feeling is that we do not, despite the fashionability of the product evaluation approach. Most of us have underestimated the difficulty of such an approach and have ignored the conceptual and measurement problems which remain to be solved. Two of the most sobering reminders of this are expressed by Travers in two papers (26; 24), one dealing with the nature of theory building, and the other with some problems of the product-oriented approach to instruction and evaluation.

In summary, it seems to me that examples of the problems which must be solved before we can begin to attempt, with any hope of success, to evaluate the graduates of programs of teacher educations are of the following classes.

1. Problems of criteria: For example, which behaviors and characteristics of teachers are going to be specified as the proposed outcomes of the program against which the graduates will be evaluated? Which characteristics and behaviors of pupils will be measured to determine teacher effectiveness?
2. Problems of criterion relevance: For example, what is the evidence that the criterion behavior specified in the outcome is relevant to

the teaching task, and has utility in facilitating learning, and is practical in the real world of teaching? With which situational and pupil characteristics does it interact?

3. Problems of measurement: For example, for which classes of teacher and pupil behavior and teacher and pupil characteristics have we reliable and valid measurement instruments and for which have we not? If we attempt measurements of natural settings (ongoing teaching-learning) as opposed to measurements of constructed "artificial" settings, how can we decrease the likelihood of sampling error?

All of these and other similar problems actually are problems for research in teacher effectiveness. The evaluation studies which are attempted can only be as good as the research basis on which they rest. And what can we say of this research basis? Biddle states unequivocally, "We do not know how to define, prepare for or measure teacher competence" (5: 13). Farther on in the same work he writes " . . . a general classification of teacher behaviors appropriate to the study of effectiveness has not been advanced--nor does it seem likely that a satisfactory system will be produced in the next decade" (5: 12). In contrast, Flanders (9), in a review based largely on his own and other related work concludes that empirical cause-effect relationships exist between certain characteristics of teachers and pupil change and that adequate instrumentation is available to permit measurement of these characteristics on a large scale. Travers (24), however, in what is, unfortunately, merely a passing reference to studies using interaction analysis, questions the extent to which we can use their results as a basis for constructing training programs.

I do not wish to belittle the import and direction of the *Recommended Standards*. Nor do I wish to discourage others here to attempt to undertake product evaluation studies. But I hope that teacher educators who may have jumped on a bandwagon will recognize that at the moment the product evaluation movement is mostly just talk and that a tremendous amount of research and development awaits us before we will have licked this problem. If this is so, I believe our strategy should be to attempt many, many, reasonably small studies each of which attempts to increase the fund of knowledge and the supply of instruments and techniques. Only in this way will we secure a better foundation for the design and evaluation of teacher education programs than presently exists.

Bibliography

1. American Educational Research Association. "Report of the Committee on the Criteria of Teacher Effectivness," *Review of Educational Research*, 22: 238-63; 1952.

2. ---. "Second Report of the Committee on the Criteria of Teacher Effectiveness," *Journal of Educational Research*, 46:441-58; 1953.

3. American Association of Colleges for Teacher Education. *Recommended Standards for Teacher Education*. Washington, D.C.: the Association, 1969.

4. Barr, A. S. *The Measurement and Prediction of Teacher Effectiveness*. Madison, Wis.: Dembar Publications, 1961.

5. Biddle, Bruce S. "The Integration of Teacher Effectiveness Research," *Contemporary Research on Teacher Effectiveness*. New York: Holt, Rinehart and Winston, 1964.

6. Corle, Clyde. *Mathematics Teaching Behavior Changes made by Intermediate Grade Teachers During a 15 Week Period of Instruction by ETV*. Final Report. University Park: Pennsylvania State University, 1967.

7. Dickson, George E., and others. *Educational Specifications for a Comprehensive Elementary Teacher Education Program*. Vol. II, *The Specifications*. Final Report. Toledo: Consortium of the State Universities of Ohio, 1968.

8. Fattu, Nicholas S. *Nine Proposals for Elementary Teacher Education*. Washington, D.C.: Department of Health, Education and Welfare, Office of Education, 1968.

9. Flanders, Ned. "Teacher Effectiveness," *Encycopedia of Educational Research*, 4th ed. New York: Macmillan, 1969.

10. Hough, John B. "An Observational System for Analysis of Classroom Instruction," AERA, 1965.

11. McGuire, Christine H., "Testing in Professional Education," *Review of Educational Research*, 38:49-60; February 1968.

12. ---. *An Evaluation Model for Professional Education--Medical Education*. Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton, N.J.: Educational Testing Service, 1968, b.

13. Mitzel, H. E. "Teacher Effectiveness," *Encyclopedia of Educational Research*, 3rd ed. New York: Macmillan, 1960.

14. Morsh, J. E., and E. W. Wilder. "Identifying the Effective Instructor: A Review of the Quantitative Studies, 1900-1952," *U.S. Air Force Personnel and Training Research Center Research Bulletin*, No AFTPRC-TR-54-44; 1954.

15. Popham, W. James. *Development of a Performance Test of Teaching Proficiency*. Final Report. Los Angeles: University of California, 1967.

16. ---, and Eva Baker. "A Performance Test of Teaching Effectiveness." Paper presented at American Educational Research Association annual meeting, 1966, Chicago.

17. Rabinowitz, W. and R. M. W. Travers. "Problems of Defining and Assessing Teacher Effectiveness," *Education Theory*, 3:212-19; 1953.

18. Ryans, David G. *Criteria-problems in Validating Teacher Selection Policies and Procedures*. Brooklyn: New York City Board of Education, June 1967.

19. ---. *Characteristics of Teachers*. Washington, D.C.: American Council on Education, 1960.

20. Sandefur, J. T., and others. *An Experimental Study of Professional Education for Secondary Teachers*. Final Report. Emporia: Kansas State Teachers College, 1967.

21. Simon, A., and E. G. Boyer, eds. *Mirrors for Behavior: An Anthology of Classroom Observation Instruments*. Vol. 1-6. Philadephia, Pa.: Research for Better Schools, 1968.

22. ---. *Mirrors for Behavior: An Anthology of Classroom Observation Instruments, Continued*. Vol. 7-15. Philadelphia, Pa.: Research for Better Schools, 1970.

23. Stiles, L. J., and R. P. Parker. "Teacher Education Programs," *Encyclopedia of Educational Research*. 4th ed. New York: Macmillan, 1969.

24. Travers, R. M. W. "Some Deficiencies of the Industrial Production Model When Applied to Education." Paper presented at American Educational Research Association annual meeting, 1968, Chicago.

25. ---. Review of *Conceptual Models in Teacher Education*, by John R. Verduin. Washington, D.C.: American Association of Colleges for Teacher Education, 1967.

26. ---. "Towards Taking the Fun Out of Building a Theory of Instruction," *Teachers College Record*, 68:[No. 1] 49-60; 1966.

27. Turner, R. L., and N. A. Fattu. "Skill in Teaching, A Reappraisal of the Concept and Strategies in Teacher Effectiveness Research," *Bulletin of the School of Education* [Indiana University, Bloomington], 36:3; 1960.

28. University of Missouri. *The Teacher Education Project of the School of Education*. Final Progress Report. Kansas City: the University 1967.

29. Woodruff, A. D. "Accreditation by Evaluation of the Product: What It Means for Teacher Education Programs," *Teacher Education: Issues and Innovations*. Twenty-first Yearbook. Washington, D.C.: American Association of Colleges for Teacher Education, 1968. pp. 237-45.

This publication was prepared pursuant to a contract with the U.S. Department of Health, Education, and Welfare, Office of Education; Contract number OEC-0-8-080490-3706-(010). Contractors undertaking such projects under Government sponsorship are encouraged to express their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.